STRUCTURE AND DYNAMICS OF HUMAN RNA G-QUADRUPLEX MOTIFS

DISSERTATION

zur Erlangung der

naturwissenschaftlichen Doktorwürde

(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

HELENA GUISET MISERACHS

aus

Spanien

Promotionskomitee:

Prof. Dr. Roland K. O. Sigel (Vorsitz und Leitung der Dissertation)

Prof. Dr. Nathan W. Luedtke

Prof. Dr. Ramón Vilar

PD Dr. Eva Freisinger

Zürich, 2016

To my parents.

To the memory of my grandmother, Aurora Serra.

To the memory of my Chemist grandfather, Francesc Miserachs.

"Almost all aspects of life are engineered at the molecular level, and without understanding molecules we can only have a very sketchy understanding of life itself." – Francis Crick, 1988

My very special thanks go to my Ph.D. supervisor

Prof. Dr. Roland K. O. Sigel

For giving me the opportunity to work on this project and for his scientific advice and support over the last four years

and to the members of my Ph.D. committee

Prof. Dr. Nathan Luedtke

and

PD Dr. Eva Freisinger

from the University of Zürich

for kindly acting as referees

and

to Prof. Dr. Ramón Vilar from Imperial College London

for the research collaboration, for welcoming me for a short scientific stay, and for his endless enthusiasm.

Acknowledgements

First and foremost, I would like to thank PROF. DR. ROLAND SIGEL for accepting me in his group and providing me with this exciting project and with many opportunities for learning and self-development. It has been an exciting ride!

Many thanks to PROF. DR. RAMÓN VILAR, PROF. DR. NATHAN LUEDTKE and PD DR. EVA FREISINGER for taking the time and interest to be part of my PhD committee.

Thank you to DR. RICHARD BÖRNER for his supervision, support and encouragement, and for always keeping the optimism alive. You really are a "catalyst", and your enthusiastic planning and confidence really made things happen!

Thank you to DR. DANIELA DONGHI for her invaluable input and advice, for her indefatigable interest and motivation, and for her patience with the (less than optimal) NMR spectra. I admire you enormously and it's really incredible how you always remember the details of everyone's research!

Thank you so much to DR. SILKE JOHANNSEN for always asking the tough questions, and for her constructive criticism that kept me pushing forward and wanting to improve.

To DR. SEBASTIAN KÖNIG, for introducing the lab to the idea of G-quadruplexes, and for the helpful discussions during my first year.

To DR. ALICIA DOMÍNGUEZ-MARTÍN, for many helpful discussions and meetings on G-quadruplexes and for being my G4 conference buddy (and shopping buddy ;)) in Grenoble and Bordeaux.

A million thanks to DR. ANITA SCHMIDTZ for always being there, from the very beginning. For being the best German teacher, for all the help, advice, protocols and tricks in the lab, and for all the personal support, inside and outside the lab. I also missed our cookie breaks when you left!

To SUSANN ZELGER-PAULUS, for being the best lab-mate: for the music, the cakes, the conversations, the biochemical tips... and for always knowing how to "get me out of my shell" on my bad days.

To PROF. DR. RAMÓN VILAR, for accepting me in his lab at Imperial College for a short, but very intense, research stay, and for always finding the time to discuss results and issues with a smile. And to DR. BEATA KLEJEVSKAJA, for being my angel during my 5-week stay in London and taking a real interest in me and my project.

To MÉLODIE HADZIC, for always being available to magically solve PC and software problems, for taking the time to advice me on Origin matters and for always asking insightful questions at the subgroup meetings. And of course for the wonderful MASH software and the hints for smFRET data interpretation!

To DR. KATSYARINA TARASAVA, DR. BHAUMIK DAVE and JOVANA JAKOVLESKA, for help with the troubleshooting of the UV spectrometers. And to ILIJA VUKADIN and DR. SOFIA GALLO for passionate discussions about the native PAGE gels.

To AUREL SPECKER, who was my student during his ACPII Praktikum, for CD measurements.

To ADRIEN MARCHAND, DR. FÉDÉRIC ROSU and DR. VALÉRIE GABELICA from the University of Bordeaux (France) for the ESI-MS measurements.

To past and present members of the smFRET AND NMR SUBGROUPS, for listening to me and discussing my results in the subgroup meetings during these past four years.

To the whole SF-LAB, for all the cake times, barbeques, apéros, birthday parties, Christmas dinners, etc., and for the good daily working atmosphere.

To DR. SERGE CHESNOV, for the MALDI-MS analysis. To DR. JAE KYONG PAK, DR. SABINE SCHTOCKHAUSE, BEATRICE SPICHTIG, NATHALIE MELUNSKI-FICHTER, and RAMONA ERNI, for the administrative support. To DR. FERDINAND WILD and MIRKO HOFER, for the IT support. To DR. THOMAS FOX and SIMON JURT for help with NMR spectroscopy, and to HANSPETER STALDER for his handywork at the mechanical workshop.

To the GRADUATE SCHOOL OF CHEMICAL AND MOLECULAR SCIENCES ZURICH (CMSZH), for financial support for conferences and for the final printing of this thesis, as well as the fantastic yearly Retreats.

To the FORSCHUNGSKREDIT OF THE UNIVERSITY OF ZURICH, for a year of funding (grant number FK-13-090).

To the kind-hearted and patient people who proof-read parts of my thesis: DR. RICHARD BÖRNER, DR. DANIELA DONGHI, SUSANN ZELGER-PAULUS, DR. MIRIAM SKILANDAT, MÉLODIE HADZIC, FLAVIEN MOREL, DR. ANITA SCHMITZ, DR. SILKE JOHANNSEN and MARIA SANZ.

To DR. ANITA SCHMIDTZ and DR. PHILIPP ÄNSTATT for the German translation and revision of the summary, respectively.

To the "ex-ACI" PEOPLE for a nice working atmosphere. Especially thank you to DR. ANASTASIA MUSIARI, DR. ERICA FIORINI, ELENA ALBERTI, MATTEO CROCE and

SIMONA CONTI for the conversations in Italian. And to ERICA again for sharing the last steps of the thesis.

To my past lab supervisors – DR. ROBERT P. DAVIS, PROF. DR. VIRTUDES MORENO, PROF. DR. PAUL DYSON, DR. JOHN J. ROSSI – and lab mentors – DR. LAURA PATEL, DR. ESTER ESCRIBANO, RUBÉN SÁEZ, DR. ALEXEY NAZAROV, DR. PRITSANA CHOMCHAN, DR. JESSICA ALLUIN and PRASANTHI DURVASULA – for their inspiration and everything they taught me, including the curiosity for research.

To DR. ANITA SCHMIDTZ, DR. MIRIAM SKILANDAT, SUSANN ZELGER-PAULUS, SERRANDA GASHI, NADIA BAATI, RIMA DRISSI, NÚRIA ALBARRÁN, MARTA GÓMEZ, CARLA GUILLÉN, DR. PHILIPP ÄNSTATT and ANDER OSES for their friendship, which made it a bit easier to be away from home.

To my friends from Barcelona, especially ANDREA ROVIRA, ALBERT FERRER, PAU ARAGONÈS, MARIA SANZ, ALBA VIDAL, SÒNIA RODRÍGUEZ, CLARA TRESERRA and MARIA TOMÉ, for always being there for me, even from the distance.

Of course, a huge THANK YOU to my parents, NÚRIA and ALBERT, for always believing in me, for supporting me through my undergraduate studies and for transmitting me an interest for science and a passion for learning.

To FLAVIEN (soon-to-be DR. MOREL), for his constant love and support, for listening to me in the good days and the bad, and for all his encouragement and calming words. And also for the tricks on format and Photoshop during the building up of this thesis. Je t'aime plus que tout mon amour!

Zürich, November 2015 H. G. M.

Table of contents

List of abbreviations

Å	angstrom(s)
ACN	acetonitrile
ALEX	alternating laser excitation
ATP	adenosine 5’-triphosphate
BB	bromophenol blue
BML	Bloom helicase
bp	base pair
BSA	bovine serum albumin
°C	degrees Celsius
calcd.	calculated
CD	circular dichroism
c-kit	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog
c-myc	v-myc avian myelocytomatosis viral oncogene homolog
Δ*ε*	molar circular dichroism
DEAD-box	family of helicase proteins containing a conserved domain with the amino acid sequence D-E-A-D (aspartate-glutamate-alanine-aspartate)
DLS	dynamic light scattering
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
ds	double-stranded
DSS	4,4-dimethyl-4-silapentane-1-sulfonic acid
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
εLCP	molar extinction coefficient for left circularly polarized light
εRCP	molar extinction coefficient for right circularly polarized light
ESI	electrospray ionization

FRET	Förster resonance energy transfer
G4	G-quadruplex
G4R1	G-quadruplex resolvase, also called RHAU (*vide infra*)
GTP	guanosine 5’-triphosphate
HPLC	high-performance liquid chromatography
htelo	human telomere
Hz	hertz
µ	micro
M	molar (moles per liter)
M^+	monovalent cation
M^{2+}	divalent cation
MALDI	matrix-assisted laser desorption ionization
MOPS	3-(*N*-morpholino)propanesulfonic acid
mRNA	messenger ribonucleic acid
MS	mass spectrometry
m/z	mass-to-charge ratio
nm	nanometer(s)
NMR	nuclear magnetic resonance
NRAS	neuroblastoma rat sarcoma viral oncogene homolog
nt	nucleotide
NTP	nucleotide triphosphate
OAc^-	acetate anion
OSS	oxygen scavenger system
PAGE	polyacrylamide gel electrophoresis
PDB	protein data bank
POT1	protection of telomeres 1
ppm	part(s) per million
py	pyridine

RecA	bacterial DNA recombination protein
RHAU	RNA helicase associated with AU-rich element
RNA	ribonucleic acid
rRNA	ribosomal ribonucleic acid
RSM	RHAU-specific motif
r.t.	room temperature
SCy3	sulfonated cyanine 3
SCy5	sulfonated cyanine 5
sm	single molecule
ss	single-stranded
TBA	tetrabutylammonium
TBA	thrombin-binding aptamer
TDS	thermal difference spectra
TEAA	triethylammonium acetate
TERRA	telomeric repeat-containing RNA
T_m	melting temperature
TPP1N	tripeptidyl peptidase 1
tRNA	transfer ribonucleic acid
UTR	untranslated region
UV	ultraviolet
v/v	volume per unit volume (volume-to-volume ratio)
vis	visible
w/v	weight per unit volume (weight-to-volume ratio)
XC	xylene cyanol

Chapter 1

Introduction

1.1 RNA

First studied in the early 1900s[1,2] and initially known as "yeast nucleic acid", ribonucleic acid (RNA) is essential for the regulation and expression of genetic information in all living organisms. From the postulation and confirmation of messenger RNA in the 1950s[3] and 1960s[4], respectively, to today's increasing knowledge on riboswitches, retrotransposons, small RNAs in post-transcriptional gene silencing and non-coding RNAs controlling epigenetics, RNA research has yielded 18 Nobel Prizes and a great number of scientific advances[5–15].

1.1.1 RNA in biology

RNA is a polymeric and polyanionic macromolecule essential for all known forms of life, in which it performs various biological roles[16–18]. According to the central dogma of biology (*Fig. 1.1.*), postulated in 1958 by Crick[3], the genomic information stored in desoxyribonucleic acid (DNA) and packed as chromosomes in the nucleus is used to synthesize the cellular actors, proteins, via the intermediate step of messenger RNA (mRNA), which serves as a template for protein synthesis in the cytoplasm[19]. Other types of RNA take part in this process. Ribosomal RNA (rRNA) forms an enzymatic complex, the ribosome, which catalyzes protein synthesis[20]. Transfer RNAs (tRNAs) carry the protein building blocks (amino acids) and recognize the mRNA sequence, using three-nucleotide codon-anticodon matches, for the correct synthesis of the corresponding protein[21].

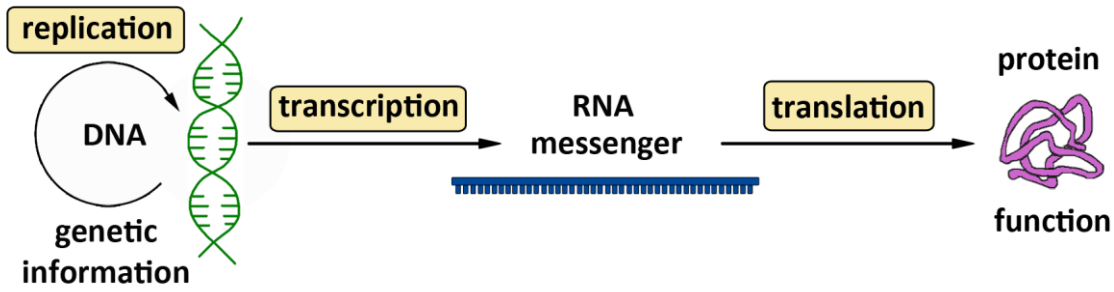


Figure 1.1. Central dogma of biology. Schematic representation of the central dogma of biology, as postulated by Francis Crick in 1958: “the transfer of information from nucleic acid to nucleic acid to protein may be possible, but transfer from protein to protein, or from protein to nucleic acid is impossible. Information means here the precise determination of sequence, either of bases in the nucleic acid or of amino acid residues in the protein”. Nowadays, this scheme is a simplified view of the RNA function in the cell (*vide infra*).

Apart from genetic information transfer, RNA is also capable of catalytic activity[22] as seen e.g. with rRNA, which does not need proteins to be enzymatically active[23,24]. Catalytic RNAs are known as ribozymes and were first reported in 1982 by Cech *et al.* with the self-splicing of the IVS (intervening sequence) in *Tetrahymena thermophila* rRNA[25]. A year later, Altman *et al.* discovered the catalytic activity of the RNA subunit in RNAse P[26].

The discovery of ribozymes and the realization of their importance in the cell lead to postulating that they might be a vestige of older life forms[27]. Indeed, one of the most widely accepted hypothesis[28,29] for the origin of life on Earth is based on an "RNA world"[30], in which self-replicating RNA molecules would have evolved from the initial inert conditions, later deriving in the last universal common ancestor (LUCA, 3.5-3.8 billion years ago containing DNA and proteins[31,32]) (*Fig. 1.2.*). The synthesis of activated pyrimidine ribonucleotides under prebiotic conditions has recently been reported and supports this hypothesis[33]. Recently, NASA scientists were able to form uracil, cytosine, and thymine from an ice sample containing pyrimidine by exposing it to ultraviolet radiation under space-like conditions[34]. Moreover, replication of short RNA sequences has been reported *in vitro* inside fatty-acid membranes[32]. The possible origin of this RNA world is still debated[35] as the appearance of ribose or other sugars from a prebiotic pool is complicated by their instability[36] and by the presence of nitrogenous substances, acting as reaction terminators[37]. An alternative possibility is that the RNA world was preceded by another replicating and evolving molecule. Threose nucleotide analogs (TNA), glycol nucleic acid (GNA), and peptide nucleic acid (PNA) have been proposed[38], as well as prebiotic monomers such as hydroxy acids, amino acids, polycyclic aromatic hydrocarbons (PAHs), and aminoaldehydes, among others[35].

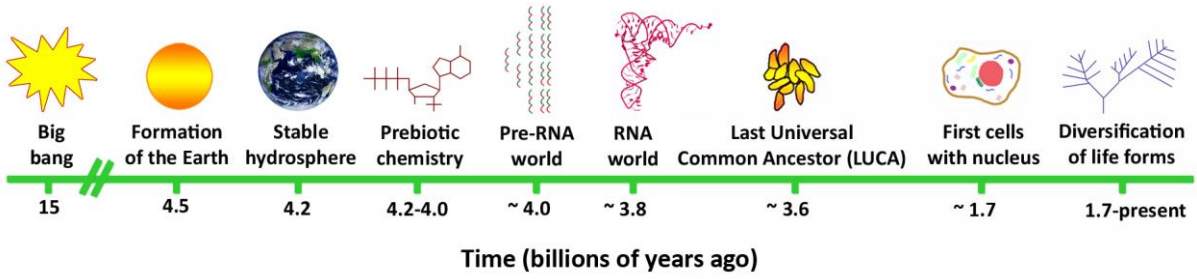


Figure 1.2. Evolution of life forms on Earth. Time scale for the evolution of life on Earth depicting the appearance of the RNA world ~ 3.8 billions of years ago, which subsequently lead to the first organisms containing proteins and, later, DNA and to the diversification of life forms into the current biosphere. Figure adapted from Joyce[39]. The RNA structure depicted as an image of the RNA world is based on the Protein Data Bank (PDB) accession code 1EHZ[40].

RNA is synthesized *in vivo* from DNA in a catalytic process known as transcription. An RNA polymerase recognizes the promoter region in the DNA template strand and opens the DNA double helix using its helicase activity. The RNA nucleotides are then linked together with release of inorganic diphosphate (pyrophosphate, PPi). With the enzyme advancing from the 3' to the 5' end on the DNA template, the nascent RNA is elongated in the 5' → 3' direction with incorporation of ribonucleotides (NTPs, *Section 1.1.2*)[19]. Post-synthetic modifications of the pre-mRNA in eukaryotes include 3' end polyadenylation (addition of a polyA tail for transport and stability)[41,42], 5' end capping (addition of a 7-methylguanylate cap, relevant for

transport, stability and translation)[43], and splicing (removal of non-coding introns and joining of the remaining coding exons)[44,45]. Subsequently and once it is mature, the mRNA is recognized by its modifications and exported to the cytosol[46], where protein synthesis takes place.

There are also RNA-dependent RNA polymerases that can use RNA instead of DNA as their template, e.g. in RNA viruses, such as the poliovirus[47]. Retroviruses, such as the human immunodeficiency virus (HIV) have the ability to transcribe RNA into DNA by using a reverse transcriptase enzyme[48]. Self-replicating stretches of eukaryotic genomes known as retrotransposons (*Table 1.1.*) use this same enzyme to move from one position to another in the genome[49]. Telomerase, an enzyme responsible for telomere elongation (*Chapter 2*) is also a reverse transcriptase carrying its own RNA template[50].

The synthesis of RNA *in vitro*, either by chemical (phosphoramidite-based synthesis)[51] or biochemical methods (transcription with viral T7 polymerase)[52], is well-established and yields oligonucleotides of defined sequence and length that can subsequently be used for *in vitro* experiments.

Non-coding RNA

Not everything in biology fits into the classical DNA-to-RNA-to-protein path (*Fig. 1.1.*). Indeed, less than half of the genomic DNA is transcribed in humans and most of the transcribed RNAs do not act as messengers for protein synthesis: ca. 98 % of the human genome is reported to be non-coding[53] (vs. coding RNA = mature mRNA). In the last decades an increasing number of non-coding RNA (ncRNA) sequences have been discovered, many of which relate to gene regulation or disease development (*Table 1.1.*), giving a new perspective and relevance to RNA presence and function in the cell.

It has been postulated that ncRNA might be responsible for the complexity and diversity of living organisms as the protein-coding DNA fraction of the genome is inversely proportional to organism complexity: ~ 90 % in prokaryotes, ~ 68 % in yeast, ~ 25 % in nematodes, ~ 17 % in insects, ~ 9 % in pufferfish, ~ 2 % in chicken, ~ 1% in mammals[54]. Moreover, in complex organisms there is an increase in alternative splicing mechanisms as well as longer untranslated regions (UTRs) in their mRNA, and the distribution of intronic sequences is nonrandom in relation to gene function[54]. These are all signs of augmented RNA signaling in more complex organisms. That is, more ncRNAs are present with regulatory functions controlling translation, mRNA half-life, chromatin architecture, transcription, splicing, RNA modification, and RNA editing[54]. Finally, it seems that ncRNAs are important in most epigenetic mechanisms (e.g., DNA methylation and histone modification regulation) and that epigenetic states can be influenced by environmental factors and inherited in higher

eukaryotes in an RNA-based, non-mendelian way[55]. Different epigenetic states allow for the appearance of different phenotypes even with the same genome, e.g. in monozygotic twins[55].

1.1.2 RNA chemical structure

The primary structure of RNA results from the attachment of monomers, called nucleotides, into a polymeric chain. An RNA monomer contains a ribose sugar with a nitrogenous heterocyclic base attached at the C1' and a phosphate bound at the C5' (*Fig. 1.3.a*). Four bases are present in RNA: two purines (adenine, A and guanine, G) binding to the pentose through their N9, and two pyrimidines (cytosine, C and uracil, U) binding through N1 (*Fig. 1.3.b*). The monomers are linked to one another via a phosphodiester bridge between the 3'-OH group of one nucleotide and the phosphate of the following. This yields a chemical directionality, 5' to 3', (*Fig. 1.3.a*) which is of high importance in all biochemical reactions involving nucleic acids.

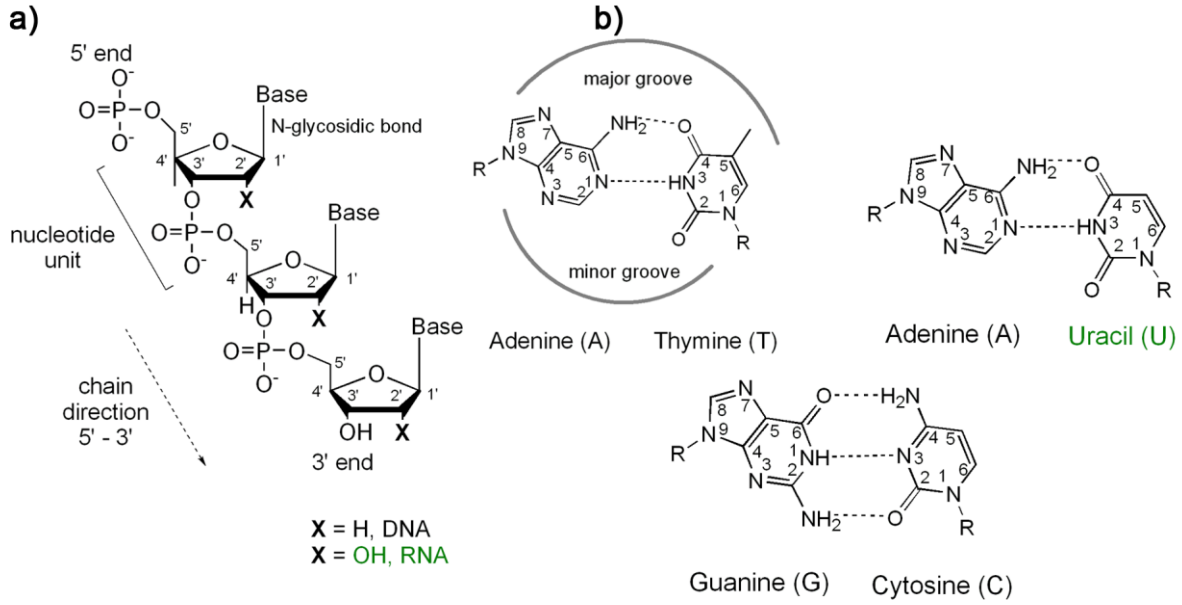


Figure 1.3. Chemical structure of nucleic acids. a) Nucleic acid chain (DNA or RNA), showing the 5' to 3' chain direction and a ribonucleotide / deoxyribonucleotide unit. b) Watson-Crick base pairs in DNA and RNA, highlighting the major and minor grooves of the corresponding double helices. Uracil in RNA is an un-methylated form of thymine. All bases are neutral at physiological pH (7.2-7.4)[56].

Table 1.1. Non-coding RNAs (ncRNAs). Non-exhaustive list of known types of ncRNAs with their average size, biological functions and non-comprehensive examples of organisms they are found in[27,57–59].

Abbreviation / full name	Size (nt)	Function	Organisms found in
miRNA / microRNA	21-25	translation regulation by mRNA degradation	*C. elegans*, *D. melanogaster*, *S. pombe*, plants, mammals, some viruses
rRNA / ribosomal RNA	~ 1500-5000	catalyzes protein synthesis; forms the ribosome	all living organisms
siRNA / small interfering RNA	21-25	(exogenous) RNA interference, gene expression regulation	*C. elegans*, *D. melanogaster*, *S. pombe*, *A. thaliana*, *O. sativa*
endo-siRNA / endogenous siRNA	21-25	mRNA and heterochromatin regulation	*D. melanogaster*, mammals
snRNA / small nuclear RNA	~ 150	spliceosomal RNAs, excise exons from pre-mRNA	eukaryotes
snoRNA / small nucleolar RNA	60-300	rRNA modification (e.g. methylation)	yeast, plants, mammals
piRNA / PIWH interacting RNA	24-31	germline development regulation	*C. elegans*, *D. melanogaster*, mammals
natRNA / natural antisense transcript siRNA	21-22	post-transcriptional gene expression regulation	*A. thaliana*
scnRNA / scan RNA	26-30	chromatin structure regulation	*T. thermophila*, *P. tetraurelia*
tncRNA / tiny non-coding RNA	22	unknown	*C. elegans*
stRNA / small temporal RNA	~ 21	post-transcriptional gene expression regulation	*C. elegans*
tRNA / transfer RNA	76-90	transfers amino acids onto polypeptide chain during protein synthesis	all living organisms
riboswitches[60,61]	40-250	located on mRNA; gene expression regulation in response to metabolites	bacteria, plants, fungi
SRP / signal recognition particle	114-300	protein recognition and transport	prokaryotes, eukaryotes
RNase P / a trans-cleaving ribozyme	140-500	tRNA processing	prokaryotes, eukaryotes
group I and II introns	200-3000	self-splicing sequences	fungi, plants, protists, bacteria
retrotransposons	100-5000	mutations by inserting near or within genes	plants, animals

The DNA and RNA bases can pair by hydrogen bonding in the canonical Watson-Crick way, which links C and G by three hydrogen bonds and A and U by two (*Fig. 1.3.b*). This results in secondary structures, e.g. the classical B-form DNA helix with the linked bases pointing inwards and the phosphates running along the outside in a spiral shape (*Fig.1.4.a and Table 1.2.*). The predominant geometry of DNA *in vivo* when it is not being replicated, transcribed, or repaired, is this right-handed B-form duplex (*Fig.1.4.a and Table 1.2.*) characterized by a wide and shallow major groove, a narrow minor groove, and a C2'-*endo* sugar pucker (*Fig. 1.5.a*).

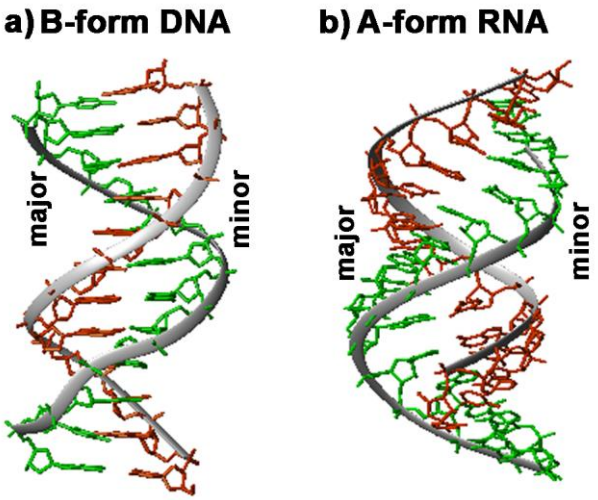


Figure 1.4. Double-stranded DNA and RNA. Helical structures showing their minor and major grooves: a) B-form DNA, b) A-form RNA. The figure was prepared from idealized B-DNA and A-RNA structures with PDB accession codes 1BNA and 1RNA, respectively[62,63].

DNA can also fold into a rarer left-handed Z-helix (*Table 1.2.*), or a right-handed A-form helix (*Fig.1.4.b and Table 1.2.*) depending on the water content and salt concentration. Meanwhile, RNA occurs in the A-form only, which is wider but more compact than B-form DNA with the nucleobases tilted with respect to the helix axis and the sugar conformation in C3'-*endo* (*Fig. 1.4.b and Table 1.2.*). The conformation of the nucleotides in the different secondary structures is also described by the glycosidic angle χ, which characterizes the relative base/sugar orientation: *syn* conformation with χ between 0 and 90°, and *anti* conformation with χ between −120 and 180° (*Fig. 1.5.b*)[64].

Table 1.2. Structural types of nucleic acids. Parameters for DNA and RNA B-helix and A-helix[65].

Structural type → / Structural parameters	B-DNA	A-DNA	Z-DNA
Rise (Å)	3.40	2.90	3.70
Twist (°)	36.7	32.7	−10/−50
Groove width (Å)	11.7/5.7	2.7/11	8.5
Helix rotation[a]	RH	RH	LH
Nº of bases per turn	10.5	11.0	12.0
Diameter (Å)	18	26	18

[a] RH = right-handed; LH = left-handed

RNA *in vivo* is more often single-stranded than DNA and rather folds intramolecularly upon itself instead of associating to a second strand. Another main difference in respect to DNA is the presence of a 2' hydroxyl, which results in extra steric hindrance and additional hydrogen bonds with surrounding water molecules, interacting ligands, or proteins, allowing a wide range of complex tertiary structures. Moreover, the 2'-OH grants RNA its characteristic instability, being easily prone to a nucleophilic attack on the adjacent 3' phosphate group, yielding self-hydrolysis of the RNA chain. This is the "price to pay" for the higher reactivity of RNA, allowing for catalytic activity. In contrast, DNA sequences are much more stable, even allowing the analysis of ancient DNA in archeological samples, recovered from animal or human remains many thousands of years old[66–68].

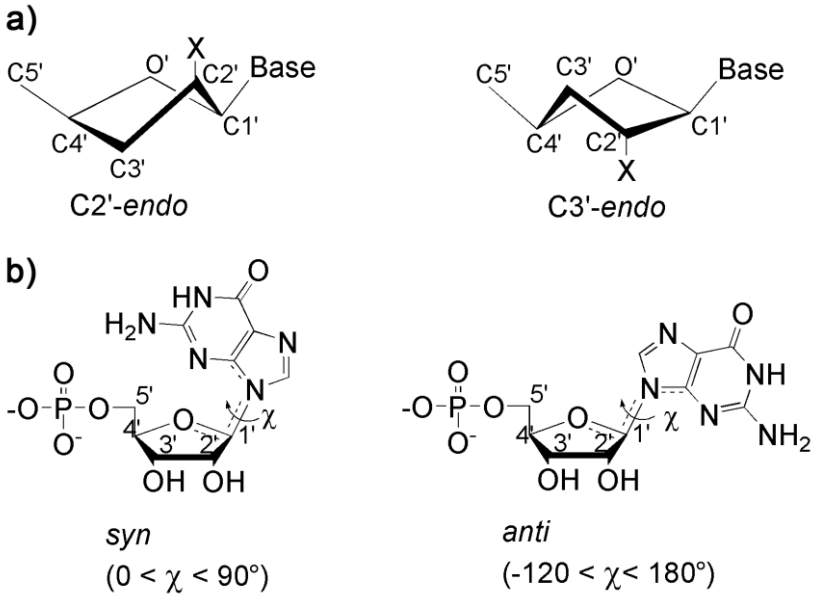


Figure 1.5. Nucleotide conformation. The conformation of the nucleotides is described by a) the sugar pucker, C2'-*endo* or C3'-*endo* and b) the glycosidic angle χ, between 0 and 90° in *syn* and −120 and 180° in *anti* [64].

Apart from the Watson-Crick base pairs (*Fig. 1.3.b*), other base associations can also form and remain stable in RNA, yielding a variety of secondary structures. G-A mismatches, G·U,

G·T and A·C wobble pairs can be formed, as well as Hoogsteen base pairs[69], where the N7 of purines acts as a hydrogen bond acceptor (*Fig. 1.6.*). These unusual base pairs are often relevant for the formation of complex three dimensional structures (*Fig. 1.7.*)[27]. Hoogsteen pairing is also present in DNA and RNA G-quadruplexes (*vide infra*)[70].

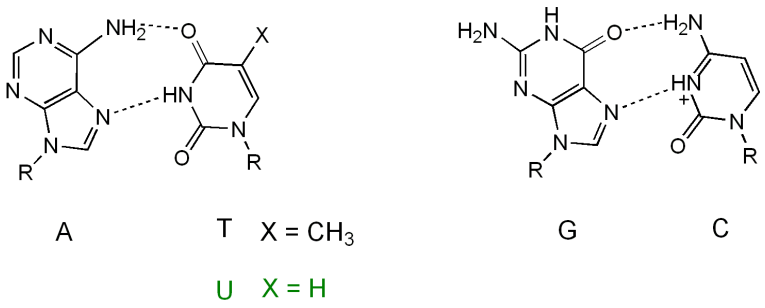


Figure 1.6. Hoogsteen base pairing in nucleic acids. Non-canonical base pairing using the Hoogsteen side of adenine (A) and guanine (G)[64].

From these diverse secondary structure elements (*Fig. 1.7.i-vii*), several tertiary structures can be formed by RNA upon folding. The simplest consist in stacking interactions, e.g. between two RNA helices, and the most usual tertiary structures are kissing loops (*Fig. 1.7.viii*) and pseudoknots (*Fig. 1.7.ix*).

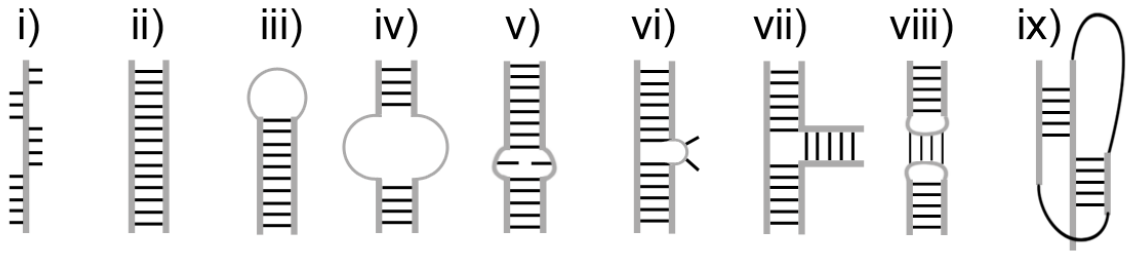


Figure 1.7. Secondary and tertiary structural motifs in RNA. Examples of RNA folding into secondary structure motifs: i) single strand, ii) double helix, iii) hairpin, iv) internal loop, v) mismatch, vi) bulge, and vii) 3-way junction. Tertiary structure motifs in RNA: viii) kissing loops, ix) pseudoknot.

1.1.3 Metal ions in RNA folding

The main driving force for RNA folding is the stacking of the bases, which increases stability of the sequence: the bases are non-polar and their association reduces the area exposed to the polar solvent (e.g., water). However, RNA is a negatively charged polyelectrolyte (or polyanion) with one negative charge present per monomer (phosphate, *Fig. 1.3.a*). Its folding therefore requires the presence of positive counter-ions to compensate the repulsion arising among the phosphate groups of the backbone and achieve electroneutrality[71].

Indeed, monovalent metal ions such as K^+ are known to promote the first step of RNA folding to secondary structures with further compaction and tertiary contacts (*i.e.*, interaction between secondary structures) being built in the presence of divalent cations such as Mg^{2+} (*Fig. 1.8.a*).

This sequential folding yields the RNA tertiary structure relevant for biological functions such as catalysis, ligand interaction, or gene regulation[72,73]. Tertiary structures are very hard to predict from the primary sequence and can be assessed by computational modeling[74] or determined experimentally by X-ray crystallography[75] and NMR spectroscopy[76].

K^+ and Mg^{2+} are the most abundant metal ions in the cytoplasm of human cells[77] (*Table 1.3.*) and are also the most relevant for RNA folding, although Na^+ and Li^+ are also known to fold some RNAs[78]. The interaction of metal ions with the RNA depends on their ionic radius, charge, charge density, dehydration energy, and coordination sphere (*Chapter 2*). The higher the charge density, the stronger the interaction with water, and the more difficult it will be to dehydrate the metal ion[71]. Being rather small, Mg^{2+} (~ 0.65 Å ionic radius) has a high charge density and is surrounded by six ordered water molecules in octahedral coordination[71]. In contrast, K^+ (1.3 Å) is accompanied by eight or nine water molecules, which are less ordered[71].

Other metal ions have also been described to perform RNA-related roles *in vivo* (*Table 1.3.*). For example, Mn^{2+} is able to replace Mg^{2+} in ribozymes and to tune their function[79]; Co^{2+} is chelated by cobalabin, which binds to B_{12} riboswitches regulating the synthesis of vitamin B_{12} in bacteria[80,81]; and Fe^{2+} can promote the folding and catalysis of RNA in the absence of oxygen[82]. Finally, protons play a capital role by defining the intracellular pH, and being involved in ribozyme acid-base catalysis.

In solution, cations accumulate around the RNA structure, divalent doing so more eagerly than monovalent, while anions are depleted from the same region[83]. Metal ions can be significantly concentrated at a distance as far as 10 Å from the RNA surface[71]. For practical purposes, three ion environments can be discerned and described as follows (*Fig. 1.8.b*)[71]:

Chelated ions

Ions that form a minimum of two direct contacts with RNA are called chelated. RNA acts as an inner sphere ligand, mainly through its non-bridging phosphate oxygens. A partial dehydration of the ion takes place[84], which depends on the ionic radius and is energetically costly. Individual free energies of binding can be calculated for each chelated ion.

Water-positioned ions

This name describes the ions which have a single layer of water molecules between them and the RNA. The electrostatic interaction that is established perturbs the hydration shells of both the ions and the RNA. Charge transfer and polarization effects are also possible in this environment.

Diffuse ions

Diffuse ions do not have any direct contact to the RNA. The hydration shell is unperturbed and the fully solvated ions are not confined to a precise location. These type of ions are a major factor in the stabilization of RNA[84], independently of the metal ion radius.

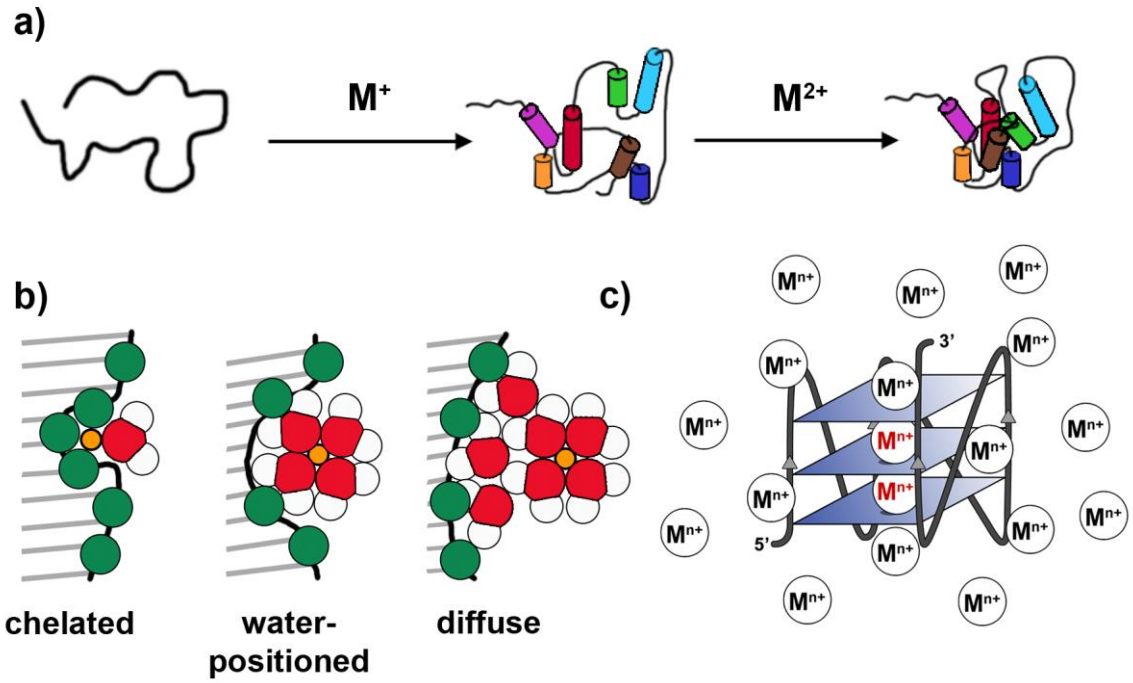


Figure 1.8. Metal ions in RNA folding. a) General pathway of RNA folding, with monovalent cations leading to the secondary structure and divalent cations allowing further compaction into the tertiary structure. b) Types of counter ion interactions in an oligonucleotide solution, according to their RNA contacts and hydration state: chelated, water-positioned and diffuse. c) Cations in G-quadruplexes can either bind specifically to the central ion channel as chelated ions (according to their radius, dehydration energy, and charge) or compensate the charges of the backbone as in other nucleotides (diffuse ions)[85].

Table 1.3. Metal ions in biology. Concentrations in human blood plasma and cells of metal ions cited herein. RNA partners and RNA-related biological roles are given, when applicable.

Cation	Conc. in blood plasma (mM)	Conc. in cells (mM)	RNA partner	Biological role with respect to RNA	Ref.
Na^+	100-200	5-15	not determined	inhibits ribosomal activity *in vitro*	77
K^+	4-5	140	all RNAs; strong stabilization of G4 RNA *in vitro*	rRNA folding *in vitro*	77
Mg^{2+}	1-2	~ 30; free: 0.5	coordinates all RNAs	ribosome stabilization; ribozyme catalysis	77
Ca^{2+}	~ 1-2	10^{-4}	not determined	inhibits group II intron splicing *in vitro*[a]	77,86
H^+	$4 \cdot 10^{-5}$	$7 \cdot 10^{-5}$	RNA bases can accept and release protons	acid-base catalysis in ribozymes	77,87,88
Mn^{2+}	$0.7\text{-}1.9 \cdot 10^{-5}$	$0.6\text{-}30 \cdot 10^{-5}$	(able to replace Mg^{2+})	tuning of ribozyme activity *in vitro*	89
Fe^{2+}	~ $2 \cdot 10^{-2}$	not determined	23S rRNA (*in vitro*)	IRPs-IRES[b] contact, mRNA; RNA folding and catalysis in absence of O_2	82,90,91
Co^{2+}	$3.4 \cdot 10^{-4}$		B_{12} riboswitches (as cobalamin)	vitamin B_{12} synthesis regulation[a]	92
Ni^{2+}	$2 \cdot 10^{-4}\text{-}1.5 \cdot 10^{-3}$	$2 \cdot 10^{-3}$	not determined	not determined	93,94
Zn^{2+}	$0.92\text{-}1.7 \cdot 10^{-2}$	$2\text{-}3 \cdot 10^{-3}$; free: $0.6\text{-}2.7 \cdot 10^{-7}$	zinc-binding domains (zinc fingers) recognize ss or dsRNA	hammerhead ribozymes cleavage *in vitro*	95,96
Li^+	$3\text{-}9 \cdot 10^{-3}$	$1\text{-}3 \cdot 10^{-2}$	not determined	not determined	97
Rb^+	$6 \cdot 10^{-3}\text{-}2 \cdot 10^{-2}$	$2\text{-}7 \cdot 10^{-2}$	(can replace K^+[a])	not determined	97
Cs^+	$2.4 \cdot 10^{-1}$ in vertebrate carcasses		(can replace K^+)	not determined	97
NH_4^+	< $3.5 \cdot 10^{-2}$	not determined	not determined	not determined	98
Sr^{2+}	~ 10^{-3}	not determined	strong stabilization of G4 RNA *in vitro*	not determined	86
Ba^{2+}	not determined	not determined	not determined	toxic, blocks K^+ ion channels	82,90-92,99-103
Cu^{2+}	$0.2\text{-}1.9 \cdot 10^{-2}$	$0.8\text{-}1.5 \cdot 10^{-2}$ (blood cells)	not determined	Cu enzymes	89,104

[a] not in humans; [b] IRP = Iron-Responsive element-binding Protein; IRES = Internal Ribosome Entry Site.

In RNA, the polynucleotide phosphates are locally fixed and in close proximity. This "electrostatic stress" from neighboring negative charges is the reason why the retention of cations is effectively larger and the exclusion of anions is smaller in comparison[84]. The number of retained monovalent ions (M^+) depends on how closely the phosphates are spaced, *i.e.* the RNA charge density. For example, an increase from 0.84 K^+/phosphate to 0.94 K^+/phosphate from ss poly(U) to ds poly(A)-poly(U)[84] can be observed. Compared to M^+, ca. half the amount of divalent (M^{2+}) ions are retained by a given RNA, e.g. ~ 0.49 Mg^{2+}/phosphate in ds poly(A)-poly(U)[84]. The fact that the same charge balancing effect is obtained with half of the particles of divalent cations represents an entropic advantage. The more folded the RNA, the higher its charge density and the greater the entropic advantage of M^{2+}. This explains why few RNA tertiary structures fold completely with monovalent cations only. In contrast, Mg^{2+} highly stabilizes the tertiary folds of RNA (*Fig. 1.8.a*).

1.2 G-quadruplexes

Since Watson and Crick published the helical structure of double-stranded DNA in 1953[105] several non-canonical nucleic acid structures have been described, such as triplexes, 3-way junctions (*Fig. 1.7.*), and quadruplexes, among others[106]. G-quadruplexes (G4s) are motifs formed in DNA and RNA when four guanine bases (G) associate through Hoogsteen hydrogen bonding forming a planar G-quartet (*Fig. 1.9.a*). Two or more quartets subsequently stack onto one another resulting in a helical conformation (*Fig. 1.9.b*)[107]. Several bioinformatics studies have shown their prevalence in the human genome[108,109] and transcriptome[110]. Interestingly, these sequences are stable under physiological conditions and are mostly found in regulatory regions of DNA[111] or RNA[112] hinting at a possible role in controlling cellular processes. Recently, DNA and RNA G4 formation has been observed in the nucleus and cytoplasm, respectively, of immobilized human cells stained with a G4-specific antibody[113,114].

As early as 1910, Bang reported that concentrated solutions of guanylic acid form a gel, which suggested that G-rich sequences may result in higher-order structures[115]. Fifty years later, in 1962, guanine self-association was described by Gellert *et al.*, who reported that guanylic acid forms layers of hydrogen-bonded tetramers[107]. Using fiber diffraction studies they proposed a planar macrocycle held together by eight hydrogen bonds and forming G-quartets staking 3.25 Å apart into a helical structure (*Fig. 1.9.a*). Shortly after, Fresco and Massoulié reported that polyriboguanylic acid forms a multistranded helix in solution and proposed also a stacking of hydrogen-bonded G-quartets[116].

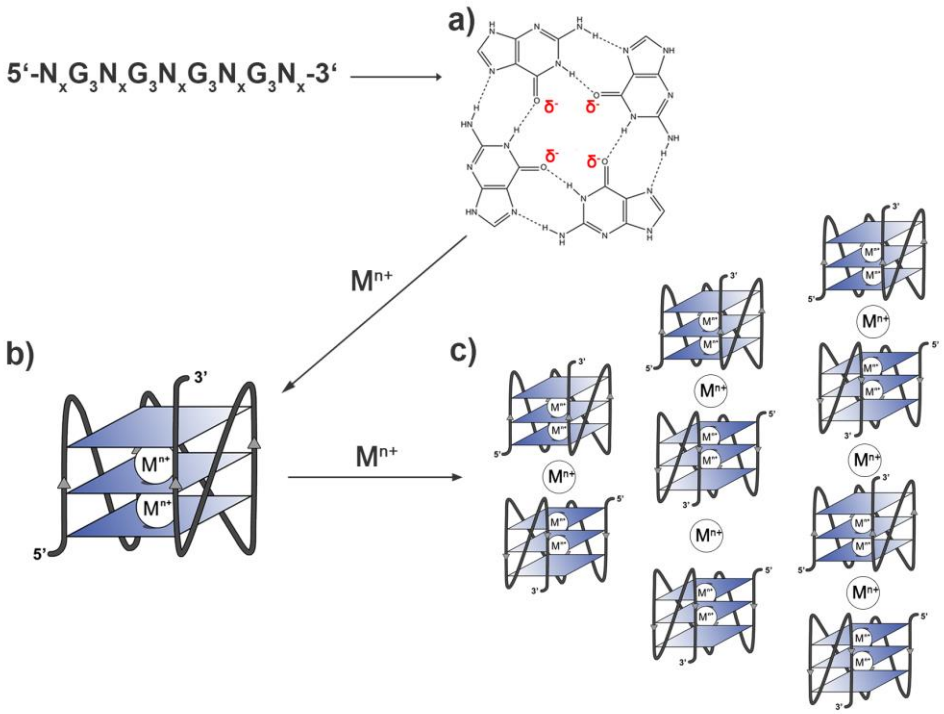


Figure 1.9. From G-rich sequences to G-quadruplexes. Starting from a G-rich nucleic acid sequence, formation of a) a G-quartet planar association of guanine bases with a central negatively-charged cavity; b) a parallel-stranded intramolecular G-quadruplex, folded by metal ion addition and π-stacking of three G-quartets; and c) dimers, trimers, and tetramers resulting from multimerization of G4 monomers in higher excess of stabilizing cations[117].

1.2.1 Structures and topologies

Guanosine has favorable properties for self-association: its edges have self-complementary hydrogen bond donors and acceptors (*Fig. 1.9.a*) and its polarizable aromatic surface bearing a strong molecular dipole is ideal for stacking. Eight hydrogen bonds are formed per G-quartet (*Fig. 1.9.a*), involving both the Watson-Crick and the Hoogsteen faces (*Figs. 1.3.b and 1.6.*). The stacked bases are polarized – the planar face is electron rich while the H edge is electron poor – and twisted around the helical axis by 30° (*Table 1.4.*). The central quartets are planar while the end quartets of a G4 motif are slightly distorted.

According to their molecularity, G-quadruplexes can be either intramolecular if all guanines are located on the same DNA or RNA strand, or intermolecular if the G-rich sequences are from different strands (*Fig. 1.10.*). Intermolecular G4s can further be described as tetramolecular, resulting of the association of four strands, or bimolecular, formed upon association of two strands (also called "hairpin dimers"). Considering the 5' and 3' ends, the G4 strands can run either in parallel or antiparallel directions giving rise to all-parallel, all-antiparallel, or hybrid conformations (*Fig. 1.10.*).

Table 1.4. G-quadruplex structural parameters. For parallel and antiparallel G4 DNAs[70].

Structural type → / Structural parameters	G4 parallel	G4 antiparallel
Rise (Å)	3.13	3.30
Twist (°)	30.0	30.0
Groove width (Å)	10.2	12
Strand polarity	++++	+-+-, +---, ++--
Helix rotation[a]	RH	RH
Nº of bases per turn	12.0	12.0
Diameter (Å)	23	21-23

[a] RH = right-handed

The linking nucleotides between the G-quartets form loops around the structure. These are not essential for folding, as it has been shown that they can be replaced by non-nucleosidic moieties[118]; but their sequence and length affect G4 stability. Depending on their lengths and type of nucleotides present the loops can have different topologies, *i.e.* lateral, diagonal, or external (*Fig. 1.10.*). Thermodynamic stability has been observed to be inversely correlated with loop size: the shorter the loops, the more stable the G4[119,120]. At the same time short loops favor the formation of parallel-stranded structures with a 4-fold symmetry. Finally, intramolecular G4s are more complex because of their additional linking nucleotides forming three loops around the structure (*Fig. 1.10.*), compared to two loops in bimolecular and no loops in tetramolecular motifs.

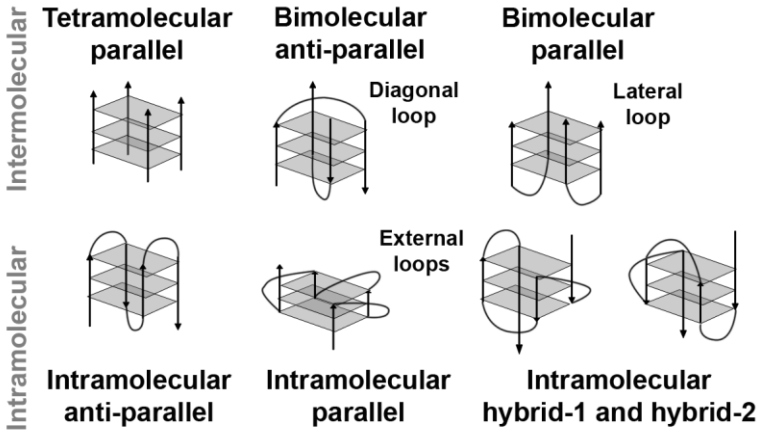


Figure 1.10. G-quadruplex topologies. Reported topologies of three-quartet G4s: tetramolecular, bimolecular, or intramolecular; parallel-stranded, antiparallel-stranded or hybrids of both. The resulting loop topologies are also shown: lateral loops, diagonal loops (requiring 2 or more nucleotides), or external loops (also called external chain reversal or propeller).

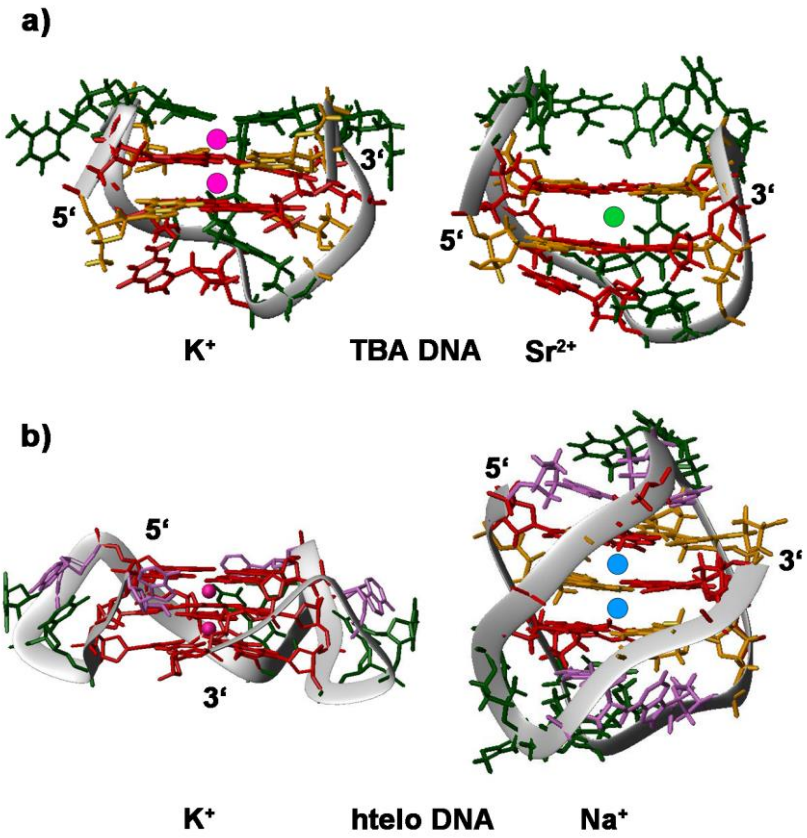


Figure 1.11. G4 DNA conformation heterogeneity. The sequence, loops, and type of cations all influence the final topology of G4 DNAs. a) Intramolecular, 2-quartet, antiparallel solution structure of the thrombin-binding aptamer (TBA) sequence 5'-GGTTGGTGTGGTTGG-3'. With K^+ (*left*), two metal ions coordinate to the structure, one in between the quartet planes, the other above. With Sr^{2+} (*right*), only one cation is present in between the quartets. The orientation of the loop bases is distinctively different in both cases. b) Intramolecular, 3-quartet structures of the telomeric sequence 5'-AGGGTTAGGGTTAGGGTTAGGG-3'. The crystal structure with K^+ (*left*) shows a parallel-stranded topology; while the NMR solution structure with Na^+ (*right*) yields an antiparallel G4. Guanines in *anti* conformation are shown in red and in *syn* in orange, thymines green, and adenines light purple. K^+ ions are magenta, Sr^{2+} ions green and Na^+ ions blue. The figure was prepared from the corresponding structures in the PDB, with accession codes a) 1C35[123] and 1RDE[124] and b) 1KF1[121] and 143D[125]. Structure 1KF1 contains metal ion coordinates from X-ray crystallography, while the remaining are determined via NMR spectroscopy and thus the cations were added artificially to show their predicted position.

The differences in linker sequences induce conformational heterogeneity as do other aspects such as nucleotide sequence, temperature, concentration, and type of cation present[70] (*Fig. 1.11.*). This makes structural predictions very challenging for G4s. Moreover, higher order stacking of G4 monomers has been observed in several NMR and crystal structures[121,122] (*vide infra*).

Both G4 DNA and RNA form either intermolecular or intramolecular motifs. G4 RNA has been shown to be more stable due to better stacking of the quartets and additional intramolecular hydrogen bonds with the 2'-OH group[126]. RNA nucleosides have a strong tendency towards

the *anti* conformation of the base in respect to the ribose sugar because of the C3'-*endo* sugar pucker, which is preferred due to steric hindrance from the 2'-OH[127]. The fact that RNA has an *anti* conformation results in G4 RNAs being always parallel-stranded[128] and thus having less conformational heterogeneity than the corresponding DNA sequences.

G-quadruplex multimerization

In vitro, short (15-25 nt) G4 sequences have been shown to aggregate by cation-mediated π-stacking of two or more G4 monomers, both in DNA and RNA. The dimeric solution structures that have been solved by NMR (*Chapter 3.1.*) show a preference for the 5' to 5' stacking orientation (*Fig.1.9.c*). It is known that sequences with short loops, which favor parallel-stranded G4s, form very stable multimers even at low nucleic acid concentrations as there is no loop impediment to G4 monomer stacking[117]. Meanwhile, G4 monomers are favored when flanking sequences to the G-repeats are added.

G4 RNAs have a higher tendency to multimerize compared to their analogous DNAs as shown by a mass spectrometry (MS) study on telomeric sequences[129]. Therefore, multimerization has to be considered during the *in vitro* studies described in this thesis as G4 RNA is always parallel and shows a high tendency towards aggregation especially when working with sequences lacking flanking nucleotides (*Chapter 3*).

1.2.2 Metal ion dependence

The presence of a cation to compensate the negative charge of the O6 oxygen atoms facing the center of the G-quartet (*Fig. 1.9.a*) is essential for formation and stabilization of G4 structures and results in a central ion channel (*Figs. 1.8.c & 1.9.b*)[130]. Metal ions that coordinate effectively inside the quartet planes increase therefore the stability of G4 motifs.

Potassium(I) is usually the most stabilizing ion in G4 DNA as it has the most suitable radius (1.33 Å) to allow coordination inside the planes as well as a dehydration energy that allows it to free itself from all coordination waters[131]. The average coordination distance has been found to be 2.73 Å with a bipyramidal antiprismatic geometry[70]. As mentioned above, K^+ is also the monovalent cation with the highest intracellular concentration (~ 140 mM)[77] (*Table 1.3.*). This further supports the probable role of G4 DNAs *in vivo*, as the most stabilizing cation is also the most common inside the cells.

A different glycosidic X angle that changes the bases from the *syn* to the *anti* conformation (*Fig. 1.5.b*) results in different stacking and different position of the ligand oxygens. Therefore, the nucleotide conformation plays a role in cation selectivity and thus it is especially noteworthy that RNA shows a high preference for the *anti* conformation[128]. The influence of

metal ions on G4 RNA structure and stability is lacking from the literature and will be the object of study in Chapter 2 of this thesis (*vide infra*).

1.2.3 Biological relevance

Recently, G4 formation has been observed in DNA in the nucleus of immobilized human fibroblasts stained with a G4-specific antibody[82]. In parallel, genome analysis showed that over 40 % of human gene promoters contain at least one putative G4 DNA[132]. The specific location and high conservation of the sequences indicates a strong selective pressure to maintain G4s and supports the hypothesis of a regulating function *in vivo*[133]. Indeed, several cellular functions have been proposed for G4 DNAs including gene regulation at both the transcriptional and translational levels, nucleosome positioning, recombination, genomic maintenance, and meiosis[134]. The first unambiguous and direct proof of G4 DNA formation *in vivo* was reported in 2001, by *in situ* immunostaining in the ciliate *Stylonychia lemnae*[135], which allowed to co-localize G4 DNA and telomeres. The telomeric regions, characterized by repeating single-stranded G-rich sequences located at the 3' end of chromosomes and acting as protecting caps for the genome of different organisms, have been the most studied G4-forming sequences to date in both DNA and RNA[136]. The human telomeric G4 DNA (htelo) is a dynamic structure, which shows different conformations in Na^+ and K^+ solution[137]. Some solution and crystal structures have been resolved[121,138–140] and investigations are currently underway concerning the telomeric proteins that might control telomeric G4 formation in the cell[141].

The presence of telomeric G4s (DNA, RNA or hybrid DNA-RNA) has an inhibitory activity on telomerase, the enzyme responsible for telomere elongation[142]. In healthy cells, the basal levels of telomerase are low and thus after several cell replication cycles the telomeres are degraded and the cells die. However, in over 80 % of human cancers there is an overexpression of telomerase, which contributes to the immortality of the tumorous cells[143]. Hence the increasing interest towards targeting telomeric G4s with stabilizing small-molecule ligands (*vide infra*) as a strategy to inhibit telomerase activity and induce tumor suppression[144].

There is also a growing interest in intramolecular G4 DNAs located in human gene promoters that are potential drug targets for which a small molecule ligand might promote G4-related transcription inhibition[111]. As an example, the G4 transcriptional repressor element in the c-myc promoter (c-myc: v-myc myelocytomatosis viral oncogene homolog) can be targeted with tetra-(*N*-methyl-4-pyridyl)porphine (TMPyP4) (*Fig. 1.13.b*)[145]. Quarfloxin entered clinical trials for the treatment of endocrine tumors by promoting c-myc G4 and inhibiting transition[146]. It showed some encouraging responses but was later discarded in phase II due to problems with bioavailability; *i.e.*, the fraction of administered dose of unchanged drug that reaches

systemic circulation[147]. Despite several proof-of-principle reports no other compounds with G4-related mechanisms have entered clinical trials so far.

RNA quadruplexes have been historically less studied than their DNA counterparts but are interestingly more likely to form *in vivo* because their native structure is single-stranded. Immobilized human fibroblast cells stained with a G4-specific antibody treated with a DNAse enzyme resulted in G4 fluorescent spots showing only in the cytoplasm[83]. These G4 loci correspond to G4 RNAs, which are known to be predominant in regulatory regions of the mRNA or ncRNA[148–150].

Bioinformatic studies have shown a high prevalence of G4-forming sequences in 5' UTRs of mRNA, which are non-coding regions containing regulatory elements[149]. An inhibitory effect of G4 on translation (*Fig. 1.12.*) has been reported for some cancer-related human sequences: neuroblastoma rat sarcoma viral oncogene homolog (NRAS)[151–153], zinc-finger protein of the cerebellum 1 (ZIC-1)[154], cyclin D3 (CCND3)[155], MT3-matrix metalloproteinase (MT3-MMP)[156], telomeric repeat-binding factor 2 (TRF2)[157], estrogen receptor 1 (ESR1)[158], and vascular endothelial growth factor (VEGF)[159]. In all these cases, formation of a stable G-quadruplex reduces the amount of protein synthesized for that particular gene by acting as a physical block for the translation via ribosome stalling. This is thought to occur either by inhibiting the assembly of the translation initiation complex or by disturbing ribosome scanning over the mRNA (*Fig. 1.12.*).

A study on eIF4A RNA helicase-dependent translation control has shown the probability of many more 5' UTR G4s regulating the transcription of oncogenes[160]. It has been demonstrated that there is a direct correlation between oncogene G4 RNA stability and the extent of translation inhibition[153]. Targeting of these G4s with stabilizing small molecule ligands has already been reported, enhancing the translational repression effect and constituting an interesting approach for novel antitumor therapies[152]. A G4 in the 5' UTR of the tumor suppressor transforming growth factor beta 2 (TGFβ2) resulted in enhanced transcription of the corresponding protein[161], which works as well against cancer development. Therefore, new druggable molecules selective for G4 regions in the 5' UTRs of either oncogenes or tumor suppressor genes might represent an attractive potential for future antitumor strategies.

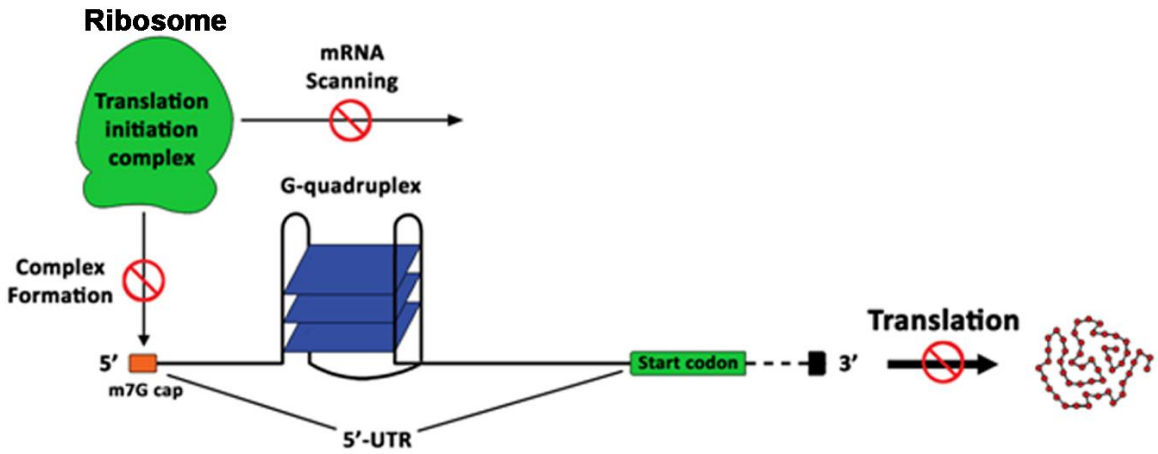


Figure 1.12. 5' untranslated region (UTR) G-quadruplex. Inhibition of protein synthesis at the post-transcriptional level by formation of a stable G-quadruplex at the 5' untranslated region of the mRNA. Translation rate decrease presumably takes place by either inhibition of formation of the translation initiation complex (containing rRNA and initiation factors) or by disturbing the ribosome scanning over the mRNA (scanning that is necessary to identify the starting codon, AUG)[148].

Further roles in post-transcriptional regulation have also been suggested for RNA G4s. For example, a splicing role where the G4 may facilitate the cleavage reaction of some human transcripts[162] or a role in the dimerization of HIV-1 RNA genome during the late state of viral replication[163]. This means that G4s might also be relevant for novel antiviral strategies as suggested by the translational repression by the viral 5' UTR G4 in Epstein-Barr virus nuclear antigen 1 (EBNA1)[164].

Among other biologically relevant applications for DNA or RNA G-quadruplexes are their use as therapeutic aptamers[165] or as diagnostic or imaging agents[150]. G4 aptamers offer several favorable properties such as non-immunogenicity, heat stability, bio-stability, ease of chemical synthesis, enhanced cellular uptake efficiency and flexibility towards the introduction of chemical modifications[150]. So far, G4-based RNA aptamers have been developed against diseases such as the Creutzfeld-Jakob encephalophathy, sleeping sickness, and Acquired Immune Deficiency Syndrome (AIDS)[150]. As some examples in diagnostic or imaging uses, a G4 RNA aptamer can detect the SARS coronavirus; and the spinach G4 RNA aptamer binds GFP and activates its fluorescence[150].

A synthetic lipophilic G4 has been shown to function as a transmembrane Na^+ transporter *in vitro*, using liposomes as membrane mimics[166]. This seems to be in accordance with the Na^+-K^+ conformational switch in some G4 sequences (e.g. htelo[167]) and with the fact that the extracellular sodium(I) concentration is high (100-200 mM) while the intracellular concentration is much lower (5-15 mM), with a reversed situation for K^+ (4 mM in blood plasma, 140 mM in the cytoplasm)[77] (*Table 1.3.*). There are also indications that the toxicity of heavy

metal ions such as Pb^{2+}, Hg^{2+}, and Cd^{2+} might be related to their binding to and disturbance of G4 structures *in vivo*[168,169].

1.2.4 Stabilizing ligands

In relation to the diverse regulatory roles postulated for G4 *in vivo*, research on small molecules that might act as G4-targeting drugs has been blooming recently[170]. Several organic molecules as well as some organometallic complexes have been demonstrated to bind G-quadruplexes and improve their stabilization. Most of the ligands discovered to date are able to selectively recognize the surface of DNA and RNA G-quartets and interact with them by π-stacking, although intercalation between the quartets and groove binding are also possible (*Fig. 1.13.a*). Well-known generic G4 binders include telomestatin, TMPyP4, BRACO-19, and pyridostatin[170] (*Fig. 1.13.b*). Very recently, organic triangulenium derivatives have been shown to interact with G-quadruplexes *in vivo*[171].

Providing ligands with enhanced selectivity towards a particular G4 topology is not an easy task and so far only modest advances have been achieved in this area[170,172–174]. For example, addition of a negatively charged functional group to pyridostatin yielded the G4 RNA-selective carboxypyridostatin[172] with the carboxyl group interacting with the 2'-OH in RNA, which is absent in DNA. This shows the importance of the 2'-OH when targeting RNA[174], as well as the involvement of the side chains in the G4 interaction. Loop involvement in ligand interaction has been described in the G4 RNA crystal structure of TERRA with an acridine ligand (*Fig. 1.14.*)[174]. The topology of the loops might be a strategy towards designing selective compounds, as e.g. diagonal loops allow the placement of bulkier ligands compared to external loops[173].

Metal complexes used as G4 ligands have the advantage over organic molecules that the metal center directs the structure via its coordination properties. At the same time, Ni^{2+} has been observed to be located on top of the K^+ ion channel in a G4 DNA crystal structure[175]. Among the tested metal complex G4 ligands are Ru(II)-, Zn(II)- or Pt(II)-complexes, Ni(II)-salphen, Mn(III)-porphyrin, and Cu(II)-terpyridine[170] (*Fig. 1.13.c*).

Currently, the available G4 ligands still lack in either specificity or affinity towards the G-quadruplex structures. For example, TMPyP4 binds to duplex, triplex, G4, and single-stranded DNA with similar affinities[176]. NMM, *N*-methyl mesoporphyrin IX, is specific for parallel G4 nucleic acids[177] but has a high binding constant, K_d ~ 2-10 µM[178]. Among the most selective ligands are so far telomestatin (*Fig. 1.13.b*) with a 70-fold binding preference for G4 vs. dsDNA[179]; NMM, specific for parallel topologies[177]; and tri-substituted isoalloxazines with up to 14-fold preference for the c-kit vs. the telomeric G4 DNA[180]. The highest reported

affinities in the concentration range of 0.2-0.3 μM[181-183] correspond to 3,4-TMPyPz, structurally related to TMPyP4 (*Fig. 1.13.b*)[181], and Cu(II)- and Pt(II)-terpyridine complexes[182].

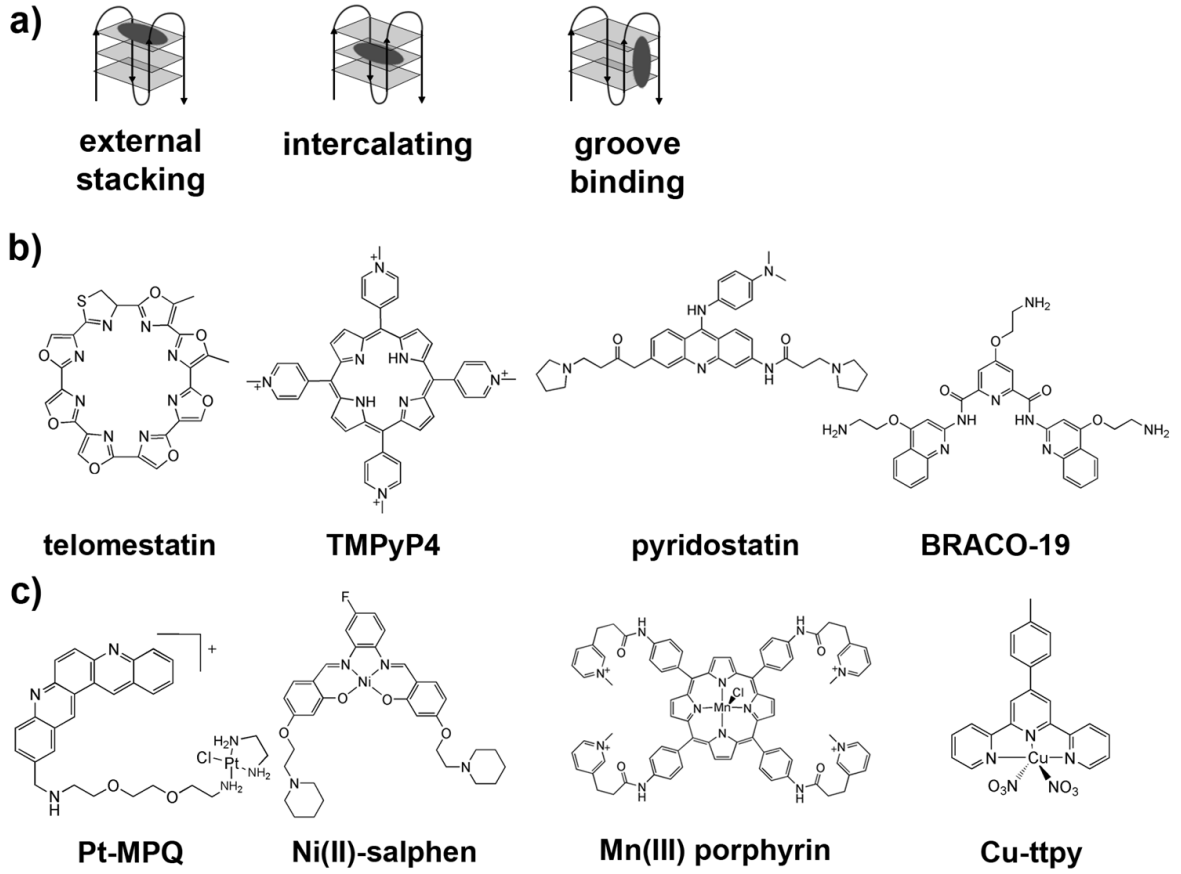


Figure 1.13. Generic G-quadruplex stabilizing ligands. a) Modes of interaction of stabilizing ligands with G-quadruplexes: external stacking onto the end quartets, intercalating between G-quartets, and groove binding. b) Organic ligands acting as generic G4 binders: telomestatin, TMPyP4, pyridostatin (PDS), and BRACO-19. c) Examples of metal complexes that interact with G4 nucleic acids: Pt-MPQ (mono-*para*-quinacridine), Ni(II)-salphen, Mn(III) porphyrin, and Cu-terpyridine[170].

Only a few reports have been published on ligands that are selective for G4 RNAs vs. G4 DNAs. Both the above-mentioned carboxypyridostatin[172] and the naphthalene diimide ligand described by Collie *et. al.*[184] contain negatively charged side chains able to interact with the 2'-OH in RNA. Other small molecules differentiate between antiparallel and parallel conformations and favour parallel G4s, and therefore, G4 RNAs. Among these are the already-mentioned NMM[177], the fluoroquinolinoacridinium cation RHPS4[185], a *N,N*-dimethylamino derivative[186] and an extended heteroaromatic 1,4-triazole[187]. However, this second class of ligands is only selective for parallel motifs but does not discriminate between RNA and DNA. Therefore, if *in vivo* targeting of G4 DNAs or RNAs for therapeutic purposes is to be achieved, further optimization of the G4 ligands will have to be performed.

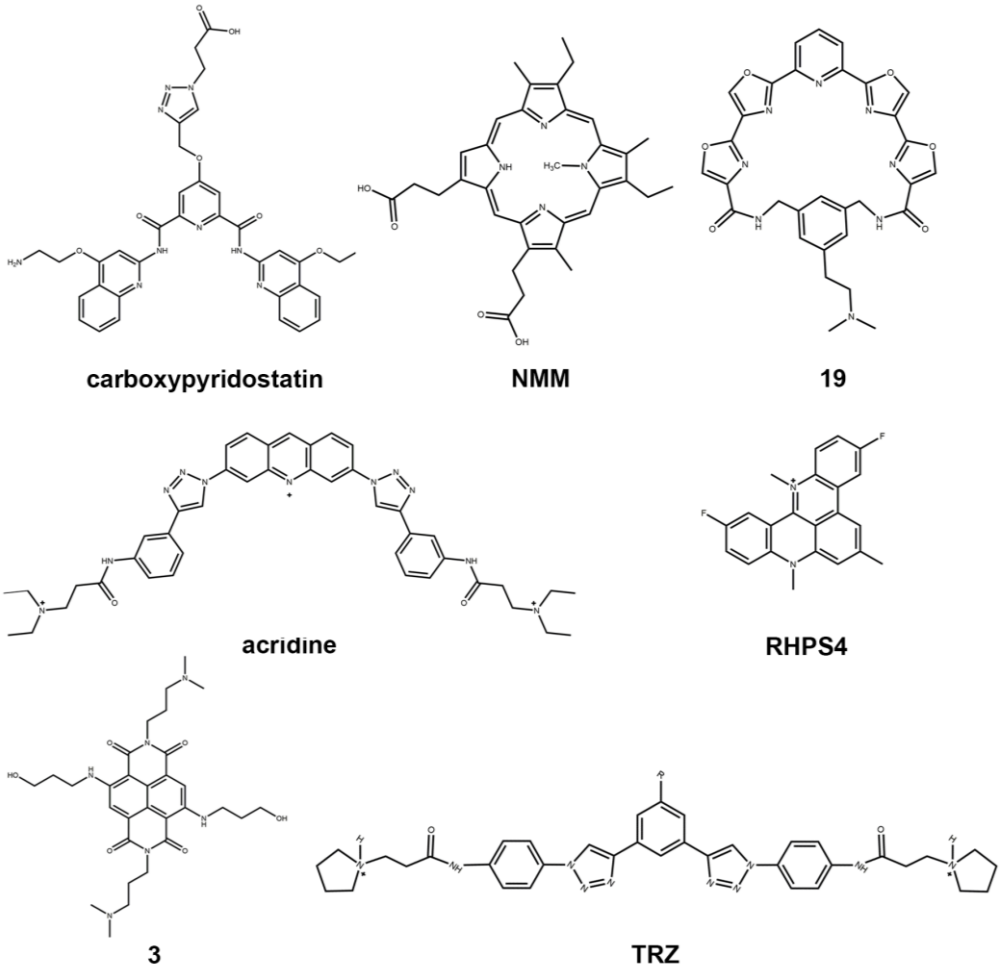


Figure 1.14. G-quadruplex stabilizing ligands selective for RNAs or parallel motifs. From top to bottom and left to right: carboxypyridostatin[172], *N*-methyl mesoporphyrin IX (NMM)[177], a *N,N*-dimethylamino derivative (**19**)[186], *N*',*N*'-((1,1'-(acridine-3,6-diyl)bis(1H-1,2,3-triazole-4,1-diyl))bis(3,1-phenylene))bis-(2-(diethylamino)acetamide)[174], the fluoroquino-linoacridinium cation RHPS4[185], a naphthalene diimide with an –OH (**3**)[184], and an extended heteroaromatic 1,4-triazole (TRZ)[187].

1.2.5 RNA G-quadruplexes under study in this project

The first G4 RNA structure was described in 1992 by Cheong and Moore, from the sequence UGGGGU and using NMR spectroscopy[188]. In this project, two G4 RNA sequences are studied by different spectroscopic and electrophoretic methods (*vide infra*) in order to elucidate their metal ion interactions, their folding, and their stability. During experiments the multimerization tendency of these motifs emerged and was therefore also analyzed.

NRAS. A highly conserved, thermodynamically stable, intramolecular G4 RNA motif (5'-GGGAGGGGCGGGUCUGGG-3', 18 nucleotides) has been reported in the 254-nt long 5' UTR mRNA of the human NRAS proto-oncogene. The NRAS (neuroblastoma RAS viral oncogene homolog) codes for a small GTPase protein (p21), which is involved both in signal transduction across the plasma membrane and in the intracellular signal pathways related to

cell proliferation and differentiation[151]. This protein is known to be overexpressed in some types of lymphomas and melanomas[189], and shows therefore a direct link with tumor progression. The G4-forming sequence is located close to the 5' cap and has been shown to inhibit translation *in vitro* in rabbit cell lysates[151]. Both the position of the G4 within the first 50 nt from the 5' end and its stability appear to be determining factors for translation repression via G4 folding[153]. The potential of the NRAS G4 as a drug target has been demonstrated *in vitro* with a pyridine-2,6-bis-quinolino-dicarboxamide containing a *p*-fluorophenyl substituent (RR110)[152].

TERRA. Telomeres, which are the regions at the ends of eukaryotic chromosomes, consist of short G-rich DNA repeats binding specialized proteins and contributing to the genomic integrity. Initially, this single-stranded DNA regions were thought to be transcriptionally silent, but later a telomeric repeat-containing RNA (TERRA) was discovered between 100-9,000 base pairs (bp) in length and forming consecutive G4 structures [188,190]. This large non-coding RNA is regulated by RNA surveillance factors in response to changes in telomere length and is thought to have roles in the regulation of telomerase and in directing chromatin remodeling throughout cell development and differentiation. We work here with a 24nt-TERRA containing four telomeric repeats and folding into one G-quadruplex, 5'-UUAGGGUUAGGGUUAGGGUUAGGG-3'. The targeting of this natural sequence with a specific ligand is proposed as a potential antitumor strategy if the telomerase enzyme activity is successfully inhibited.

1.3 G-quadruplex characterization methods

Many experimental techniques are used to study G4 formation under different conditions, *i.e.* varying DNA/RNA concentration, pH, temperature, cation type and concentration, and buffer. X-ray crystallography and NMR spectroscopy allow structure determination with atomic resolution, but several other methods give valuable information on stability and conformational changes. The techniques used in this thesis are briefly described in this section.

1.3.1 Thermal melting

The melting of G4-folded DNA or RNA can be monitored by a hypochromic shift resulting from the unstacking of the bases and the unfolding of the structure[70]. The spectral absorbance is measured at 295 nm as a function of temperature[191] using a UV-visible (UV-vis.) spectrophotometer. At this wavelength the denaturation of the sample yields a significant change in absorbance of ca. 50 %, which is easier to follow in contrast to the small absorption difference that G4s yield at 260 nm, the typical wavelength for the monitoring of duplex DNA unfolding. The melting profile obtained for a G4 is inverted as compared to the one obtained

for a DNA/RNA duplex (hypochromic vs. hyperchromic shift), allowing to easily recognize G4 formation (*Fig. 1.14.*).

A two-state model can be used to fit the UV data and obtain the melting temperature[192] (T_m), which is indicative of the DNA or RNA structural stability. As mentioned above (*Section 1.2.3.*), the stability of G4s depends on the type and concentration of metal ions present in solution. Thermal melting experiments allow determining the relative stability of the same G4 under different cation conditions. Thermodynamic parameters can easily be calculated from the melting curves if the system fulfills the two-state hypothesis, meaning that only one folded and one unfolded state are interconverting[192]. If intermediate conformations or partially folded species are present, accurate thermodynamic parameters are complicated to obtain, as it is also the case when multimerization phenomena occur[187]. Finally, the molecularity of the structures can also be assessed because formation of intramolecular motifs yields a T_m that is independent of the nucleic acid concentration. For all G4s, the melting temperature is independent of the pH in the 5-7.5 range[191].

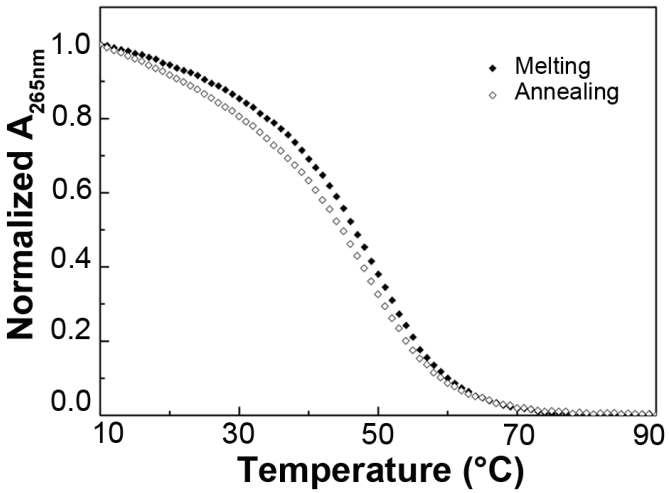


Figure 1.15. UV thermal melting curves. TERRA absorbance curves with 10 mM Na^+ (melting, 10 to 90 °C; annealing, 90 to 10 °C) in 10 mM lithium(I) MOPS buffer, pH 7.40. The G4 RNA concentration was 4 µM and the calculated T_m, 49±1 °C.

1.3.2 Circular dichroism (CD)

Circular dichroism spectroscopy (CD) is a qualitative method for the characterization of the G4 folding topology. It is based on the differential absorption by chiral molecules of left and right-handed circularly polarized light[193]. The ellipticity of the sample is recorded at room temperature *versus* the wavelength in the 230-to-350-nm range[194]. CD spectra of nucleic acids are highly sensitive to changes in nitrogen base stacking. Therefore, characteristic spectral signatures can be recognized for a parallel or antiparallel G4[70], with parallel G4 having a positive ellipticity maximum at 264-265 nm and a negative minimum at 240 nm, while

antiparallel G4 yields a positive ellipticity maximum at 295 nm and a negative minimum at 265 nm (*Fig. 1.15.*).

G4 DNA can form either the parallel or antiparallel topology, or either a hybrid fold, which yields a combination of the two signatures. In contrast, G4 RNA has been observed so far to form only in a parallel fashion[128]. Both G4 folds can be distinguished from unstructured DNA or RNA, which has a positive maximum around 270 nm, and from duplex DNA, with positive signal/s around 260-280 nm and a negative band around 245 nm (the exact position and amplitudes of the bands highly differ according to the B-DNA sequence). A-form duplexes yield a very similar spectrum to parallel G4s, which can still be distinguished from the positive sign of the signal appearing at 210 nm, which is negative in the A-duplex[195]. Moreover, the higher the intensity of the CD signal, the more DNA or RNA in the sample is folded into that particular topology. Recording thermal melting experiments is also possible with CD by following the ellipticity as a function of temperature.

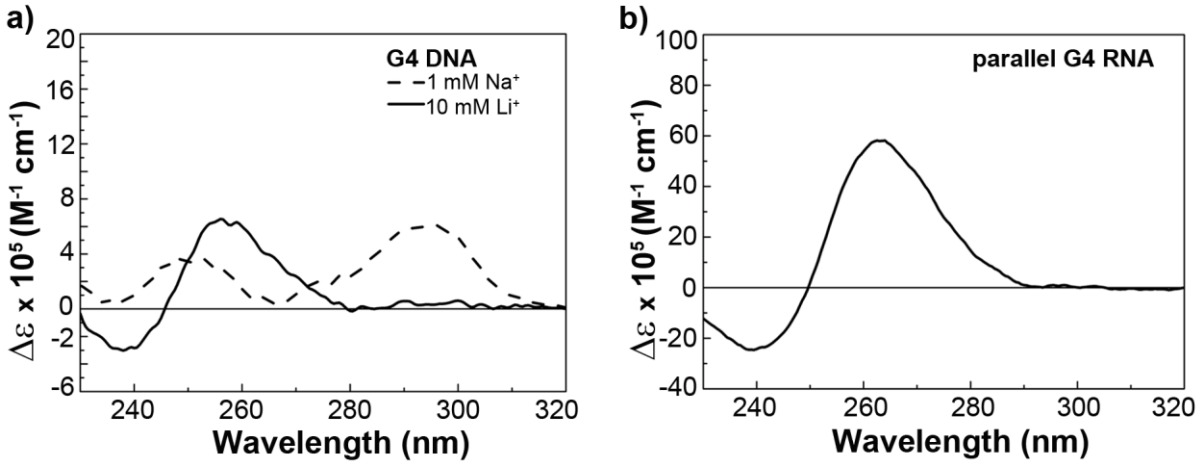


Figure 1.16. G-quadruplex topology in circular dichroism. Circular dichroism spectra showing the characteristic signals for a G4 DNA. a) antiparallel htelo DNA in 1 mM Na^+ and parallel htelo DNA in 10 mM Li^+, both at 10 µM DNA in MOPS buffer, pH 7.4; b) parallel-stranded NRAS RNA, 10 µM in 10 mM lithium(I) MOPS, pH 7.4.

1.3.3 Thermal difference spectra (TDS)

The name thermal difference spectrum designates the mathematical subtraction of a UV spectrum recorded at high temperature (above the T_m) and a UV spectrum measured at low temperature (below the T_m). This difference is expressed as Δ*Absorvance* (Δ*A*) vs. *wavelength*, and gives a characteristic shape for each type of DNA or RNA structure[196].

Although the exact position and intensity of the signals depends on the type of monovalent cation and on the exact nucleotide sequence, the global shape of the TDS is specific for each DNA or RNA structure. Therefore, G-quadruplexes can be recognized by the following signature (*Fig. 1.16.*):

- one positive band at 243±2 nm
- one positive band at 273±1 nm corresponding to the temperature-dependent increase in absorbance of the single nucleotides
- one negative band at 295±1 nm, which may vary significantly in intensity and is also seen in other DNA motifs

G4 RNAs yield related but distinct spectra in comparison to G4 DNAs[196].

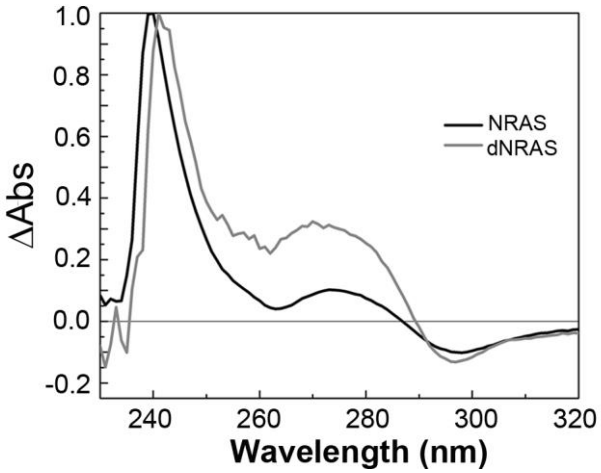


Figure 1.17. Thermal difference spectra of the G4 RNA NRAS and its corresponding G4 DNA dNRAS. UV spectra were measured at 20 °C and 90 °C in 10 mM lithium(I) MOPS with 20 mM K^+. The TDS signatures have been normalized (ΔAbs = 1 for the highest positive signal) and are therefore concentration independent.

1.3.4 Polyacrylamide gel electrophoresis (PAGE)

Polyacrylamide gel electrophoresis (PAGE) can be used for DNA or RNA detection, quantification, purification by size, and quality assessment. Nucleic acids are negatively charged and migrate toward the anode in the presence of electric current. This migration is in proportion to their mass / chain length[197] and can also be affected by topology.

Denaturing PAGE gels can resolve DNAs or RNAs from 20 to 600 nt and are run at room temperature in the presence of a denaturing agent, e.g. urea or formamide. Native gels are used to resolve different conformers, or ligand-nucleic acid complexes. In contrast, native conditions allow the nucleic acid to maintain its folded structure and consist in running the gel at low temperature (4 °C) and in presence of cations contained in the gel, the running buffer, and the samples. Visualization can subsequently be carried out either by UV shadowing, or by dye staining followed by imaging with a fluorescence scanner.

Gel studies allow to see G4-ligand interactions, to differentiate folded from unfolded RNA, to observe whether more than one folded conformation is present, and to assess the effect of metal ions. The cation concentration will determine whether or not any G4 aggregation is present in the samples. Native PAGE gels can therefore be used to elucidate the multimerization state of G4 DNAs or RNAs. For example, the monomer to dimer transition of the parallel-stranded G4 DNA d(AGGGGGAGGGAGGGTGG) is observed in PAGE gels between 0.1 and 3 mM K^+ [198] (*Fig. 1.17.*) by a shift in the sample mobility.

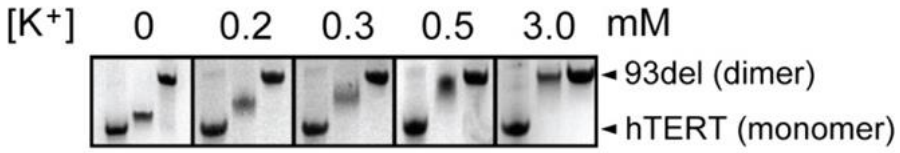


Figure 1.18. Native polyacrylamide gel electrophoresis. Images of CEB1 G4 migrated in native PAGE at different K^+ concentrations and detected by UV shadowing. hTERT and 93del were used as references for monomeric and dimeric G4s, respectively. *Reprinted with permission from ref.* [198]*. Copyright (2014) American Chemical Society.*

1.3.5 Electrospray Ionization Mass Spectrometry (ESI-MS)

Electrospray Ionization Mass Spectrometry (ESI-MS) was first used to detect G4 formation in 1993[199]. Since then, this technique has been employed to elucidate the strand stoichiometry[200], the number of ions in the central channel[201], and the ligand interactions in G4s[201,202].

In a mass spectrometer, the sample is ionized into charged molecules or molecule fragments and the mass-to-charge ratios (m/z) are measured. The electrospray ionization method uses soft conditions, *i.e.* capillary temperature and acceleration voltage both as low as possible, in order to prevent the destruction of the G4 structure during ionization[203]. Preferably the negative ion mode is used, as the nucleic acid backbone is negatively charged in solution[203].

The measured m/z allows to determine the molecular weight of the ionized G4 and the number of cations inside its central channel. Subsequently the number of G-quartet planes can be predicted knowing the number of ions that are located inside the structure[129] (*Fig. 1.18.*).

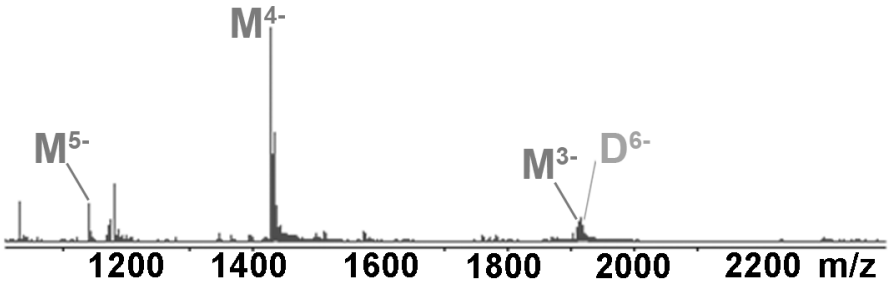


Figure 1.19. Electrospray mass spectrometry of G4 RNA. Electrospray mass spectrum of annealed dNRAS DNA at 5 µM in 100 mM NH_4OAc. The peak annotations $[M/D]^{z-}$ indicate the strand stoichiometry, monomer (M) or dimer (D), and the total charge (z). Spectrum recorded at the University of Bordeaux by A. Marchand. MW calcd. 5700.59 g/mol.

Sodium(I) and potassium(I) ions have to be avoided in ESI-MS because they condense around the negatively charged nucleic acid backbone during solvent droplet evaporation. This results in additional [DNA/RNA + n Na^+ / K^+] adducts that are heterogeneous in mass and lead to a loss of sensitivity and mass accuracy. Therefore, NH_4^+ conditions are usually used instead. With ammonium, the ions that are not coordinated between G-tetrads evaporate as NH_3 and thus salt adducts are avoided[203].

1.3.6 Dynamic light scattering (DLS)

The resulting scattering pattern from shining light on small particles in solution allows the measurement of their geometrical structure and state of motion. Dynamic light scattering2 (DLS) measures the diffusivity of small particles undergoing Brownian motion[204] and provides a measure of the hydrodynamic radius, r_H. To chose the appropriate model to calculate the expected radius and compare it to the measured values, the ratio *q* of the theoretical length *L* of the nucleic acid and its diameter *d* (*q* = *L*/*d*) needs to be computed. The theoretical length *L* can be estimated assuming an average distance of 0.313 nm between individual bases in a parallel quadruplex[70]. For the diameter *d*, a value of 2.4 nm typical for a parallel G4 structure is used (measured in two G4 PDB structures, 3IBK and 1K8P[121,282]). If *q* > 2, a spherical model can be applied to calculate the theoretical hydrodynamic radius[205]:

$$r_H = \frac{L}{2} \qquad (1.1.)$$

If the length exceeds the RNA diameter, *i.e. q* > 2, then the molecule cannot be considered spherical and a symmetrical cylinder model is used instead[205]:

$$r_H = \frac{L}{2(\ln q + 0.312 + 0.565\, q^{-1} - 0.1 q^{-2})} \qquad (1.2.)$$

1.3.7 Nuclear Magnetic Resonance (NMR)

Nuclear Magnetic Resonance (NMR) spectroscopy in solution can be used to obtain G-quadruplex structures with atomic resolution and to study their dynamic and kinetic properties, as well as their ligand interaction[206]. The formation of a G-quartet gives rise to characteristic guanine imino protons (H1), which can be observed within the 10-12 ppm chemical shift range[206] (*Fig. 1.19.*). This is in contrast to the imino signals for Watson-Crick base pairing, which appear at 13-14 ppm. The number of observed imino proton resonances correlates with the number of guanine bases involved in G-quartet formation.

Moreover, another characteristic of G4s is that their guanine imino protons exchange more slowly with the solvent. This is especially true for the imino protons of guanines in the central G-quartet, which can still be detected long after the sample is dissolved in D_2O. Therefore, ^{1}H NMR spectroscopy can be used to identify G4 formation and to observe whether different topologies are present. Indeed, the intensity of a peak is a measure of the molar concentration of the associated proton in that particular conformation[206].

Another NMR spectroscopy method, diffusion ordered spectroscopy (DOSY)[207], can be used to evaluate different G4 topologies as well as the stacking or multimerization of G4 motifs[208]. It is based on the principle that the translational diffusion of a G4 molecule in solution reflects its size and shape.

Interaction of G4 structures with cations can also be observed with NMR spectroscopy. The effect of the metal ion can be assessed on the ^{1}H spectra and a ^{15}N NMR spectra can be used with G4-NH_4^+ solutions to detect the position and movement of ammonium ions in the central channel[182, 209,210]. ^{39}K or ^{23}Na NMR spectroscopies have also been reported with G4 solutions containing either K^+ or Na^+, respectively[211-213].

Finally, NMR spectroscopy allows to study the effect of either crowding agents or ligand binding on the G4 motifs. Crowding agents, e.g. polyethylene glycol (PEG), acetonitrile (ACN), dimethyl sulfoxide (DMSO), or ethanol, have been shown to induce conformational transitions in htelo G4 DNA[214]. Indeed, the water depletion effect that they trigger favors the parallel G4 fold in detriment of the antiparallel structure[215]. Therefore, these agents can be used as a strategy to reduce conformational hetereogeneity[215]. In addition, ligand binding can be followed by titration of the ligand into the G4 sample, which often results in an observable broadening and/or shift of some G4 protons next to the binding site[216].

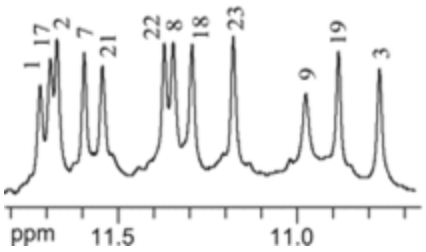


Figure 1.20. Imino protons in G-quadruplexes. Imino proton region of the 1D ^{1}H NMR spectrum of bcl2Mid G4 DNA (5'-GGGCGCGGGAGGAA*TT*GGGCGGG-3'), showing the assigned protons over each resonance. Conditions: 25 °C, 20 mM potassium(I) phosphate, 40 mM KCl, pH 7.0, 1.5 mM DNA. *Reprinted with permission from ref.* [217]*. Copyright (2006) Oxford University Press.*

1.3.8 Single-molecule Förster Resonance Energy Transfer (FRET)

First reported for DNA in 1996[218], single-molecule Förster resonance energy transfer (smFRET) allows the observation in real time of conformational dynamics of single molecules by tracking fluorescence changes[219]. This single-molecule technique has been increasingly used over the last two decades to answer biological questions[220-222]. FRET was first reported by Förster in 1948[219]. It takes place between a donor (D) and an acceptor (A) fluorophore via dipole-dipole interaction (*Fig. 1.20.*) if they are close enough so that the emission of the donor is transferred non-radiatively to the acceptor. This typically happens within a distance *R* of 2-10 nm[220]. Closer than that, a quenching effect between the dyes disturbs FRET. The transferred energy is calculated as follows:

$$E_{\mathrm{FRET}} = \frac{1}{\left(1 + \frac{R}{R_0}\right)^6} \qquad (1.3.)$$

where R_0 is the Förster radius of the fluorophore pair, defined as the distance where the FRET efficiency (E_{FRET}) is 50 % (*Fig.1.22.*).

Fluorescent dyes and RNA labeling

An ideal fluorophore or fluorescent dye has to be bright, *i.e.* with an extinction coefficient ε > 50'000 M^{-1} cm^{-1} and a quantum yield QY > 0.1. Additionally, it should be photostable, relatively small (< 1 nm, so that it does not perturb the RNA structure), water-soluble, and commercially available in a reactive form that can be bio-conjugated to the RNA[219]. It should show little intensity fluctuations and should be excitable and emitting in the visible light range[224].

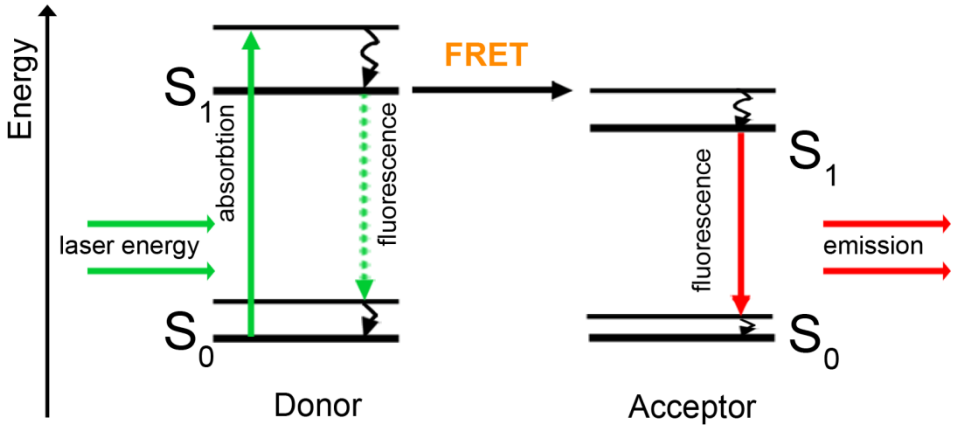


Figure 1.21. Förster Renonance Energy Transfer. Jablonski diagram illustrating the principle of FRET. After absorption of light, the donor is excited to state S_1. Vibrational relaxation brings the donor back to a lower energy state, from which energy transfer to the excited state of the acceptor can occur via a dipole-dipole interaction. Subsequently, the acceptor releases measurable fluorescence by returning to its singlet ground state, S_0.

An optimal FRET pair has the following characteristics: large spectral separation between the donor and acceptor emissions, similar quantum yields and detection efficiencies for both donor and acceptor[219], and an overlap between the donor emission and the acceptor absorption[224] (*Fig. 1.21.*).

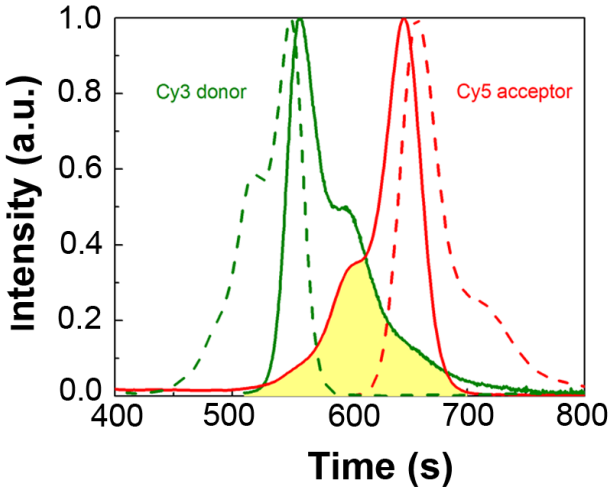


Figure 1.22. Cyanine dye pair for FRET. Overlay of normalized absorption (dashed green line, full red line) and emission (full green line, dashed red line) spectra of Cy3 (green) and Cy5 (red). The spectral overlap between the Cy3 emission and the Cy5 absorption (solid lines) allows for FRET to occur and is filled in yellow.

The most used FRET pair in nucleic acid studies is Cy3-Cy5, with two cyanine (Cy) dyes (*Fig. 1.22.a*). Cy3 acts as the donor and Cy5, as the acceptor. They both have ε over 10^5 M^{-1} cm^{-1}, and comparable quantum yields of approx. 0.15-0.30 for Cy3 and 0.3-0.4 for Cy5[220,224]. The spectral separation is ~ 100 nm[224] and they are both commercially available as amino-,

thiol- and NHS-ester derivatives, among other reactive forms. Sulfonated versions of both dyes (*Fig.1.22.a*), which are negatively charged, are used in this work in order to increase the solubility in water and minimize unwanted interactions with the polyanionic RNA[225].

To obtain the maximal sensitivity with smFRET the labeling sites have to be selected in order that the inter-dye distance is in the range of the Förster radius (*Fig. 1.22.b*) with the expected RNA dynamics. This will yield a maximum contrast in FRET efficiency between different RNA folded states.

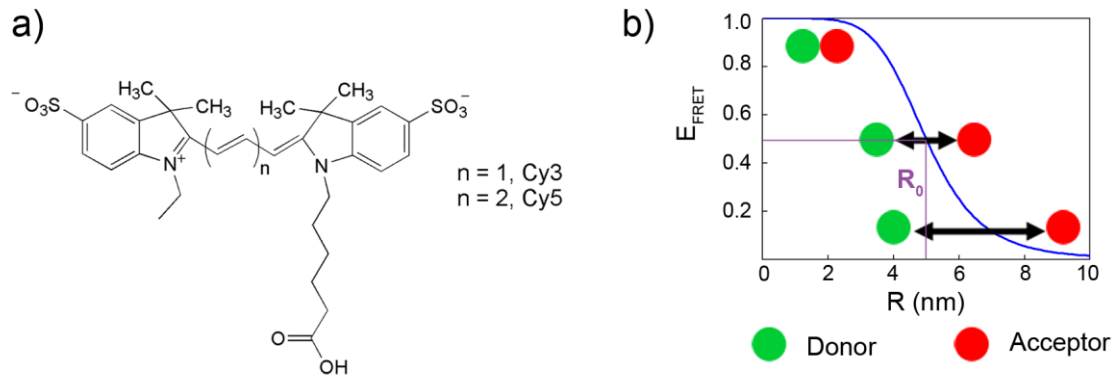


Figure 1.23. Cy3 and Cy5 dyes. a) Cyanine dyes commonly used in FRET and chosen as well for this project. b) Dependence of the FRET efficiency (E_{FRET}) on the inter-dye distance R. R_0 ~ 60 Å for the Cy3-Cy5 pair[224].

smFRET and advantages of single-molecule detection

We use a total internal reflection fluorescence (TIRF) microscope, in which an evanescent wave penetrates only 100-200 nm from the quartz surface into the solution, reducing the background dramatically[219]. The fluorescence signal exiting the sample is collected with a water immersion objective (*Fig. 1.23.*). A long pass filter is subsequently used to discard the scattered light from the excitation source before the donor and acceptor emissions are split with a dichroic mirror and placed side by side on the final recorded image. The detection is carried out by an EM-CCD camera (electron-multiplying charge-coupled device) with a high quantum efficiency (> 0.85) at 450-700 nm. This camera has a low thermal noise at −80 °C, a frame rate of 10-20 Hz, and allows imaging of hundreds of molecules in parallel by the recording of real-time movies.

It is important to consider that the presence of molecular oxygen originates highly reactive oxygen species that cause photobleaching of the dyes[219]. Therefore, an oxygen scavenger system is added to the imaging buffer. It contains glucose oxidase, catalase, and D-glucose[226], which undergo the following reactions:

i)

$$\text{D-glucose} + O_2 + H_2O \xrightarrow{\text{glucose oxidase}} \text{gluconic acid} + H_2O_2$$

ii)

$$H_2O_2 \xrightarrow{\text{catalase}} H_2O + 1/2\ O_2$$

The gluconic acid obtained as a by-product reduces the pH, and therefore a buffer is needed to maintain a constant pH value. At the same time, oxygen removal results in the fluorophores spending more time in their triplet dark state (a non-emitting electronic conformation), a fact that derives in fluorescent intermittency, also called blinking. To avoid this, Trolox (a vitamin E analog) is added to the imaging buffer and functions as a triplet-state quencher[219].

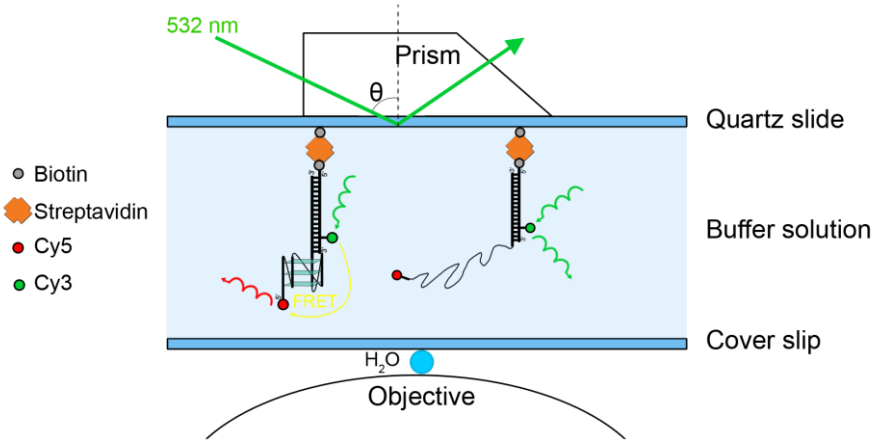


Figure 1.24. Prism-based TIRF setup for smFRET. Surface immobilization onto quartz slides via biotin-streptavidin linkage[219]. BSA, streptavidin, and quartz are negatively charged at neutral pH, repelling the RNA and avoiding unwanted surface interactions[227].

The main advantage of single-molecule detection is that there is no time nor population averaging, allowing the identification of rare or short-lived states in heterogeneous systems (*Fig. 1.24.*)[224]. Moreover, immobilization onto a surface (*Fig. 1.23.*) permits continuous tracing of the same molecules, from 1 ms to minutes in camera-based detection with TIRF illumination. smFRET directly reveals the timescales of each folding state, as well as the presence of transient intermediates[228]. It acts thus as a "molecular ruler", giving a good estimation of the distances and compaction states along a folding pathway[222].

There are, however, limitations to this technique, namely that it is insensitive to distance changes outside a certain inter-dye distance range (e.g. 2-8 nm distance for R_0 = 5 nm), that the time resolution is limited by the frame rate of the camera (1-10 ms at its lowest limit), and that the absolute distance estimation is challenging, because it needs to take the dye orientation into account, among other aspects[219]. Finally, tracking in live cells has been described but is still challenging[229,230].

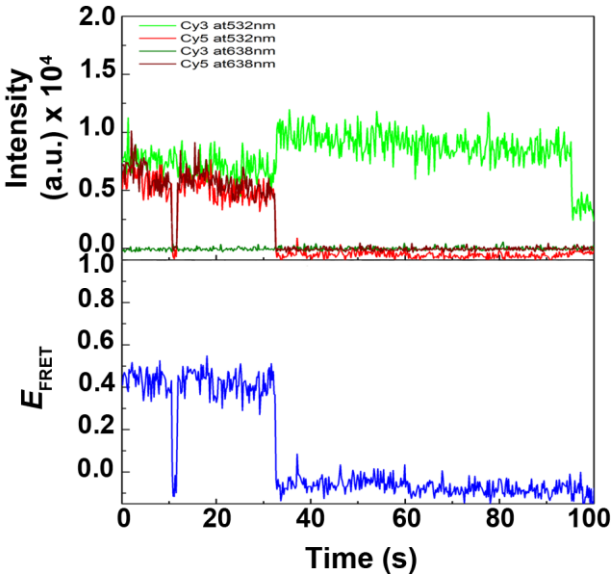


Figure 1.25. Data obtained from smFRET measurements. Example of a time trace showing the intensities of donor and acceptor (top), from which the apparent FRET efficiency (bottom) is calculated. The anticorrelation between donor and acceptor signals is typical of FRET. The Cy5 shows transitions to dark states, *i.e.* photoblinking (~ 10 s, transient) or photobleaching (~ 30 s, permanent).

1.4 Thesis outline and aims

The field of G-quadruplex structures has been thriving in the last decades and is currently a hot topic in molecular and structural biology. As it becomes clear that these motifs most likely have important regulating functions *in vivo*, improvement of the selectivity and potency of small molecule ligands might lead to the design and implementation of novel anticancer strategies.

Less is known about RNA G4s compared to their DNA counterparts, but the presence of these structures in the 5' UTR mRNA of proto-oncogenes, located in the cytoplasm of cells, is promising for the targeting with small molecule drugs. Indeed, only very recently the presence of G4 RNAs outside the nucleus was confirmed by staining immobilized human fibroblasts with a G4-specific antibody[114].

Taking advantage of the Sigel lab expertise on RNA work and RNA *in vitro* transcription, this project sets up to handle two G4-forming RNA sequences: the telomeric RNA (TERRA) and a 5' UTR, oncogenic G4 (NRAS). The main aims of the thesis are the following:

1. Observe via spectroscopic techniques the effect of several monovalent (Li^+, Na^+, K^+, Rb^+, Cs^+, and NH_4^+) and divalent (Mg^{2+}, Ca^{2+}, Sr^{2+}, and Ba^{2+}) metal ions on the folding and stability of two highly conserved, biologically relevant, RNA quadruplex-forming sequences: NRAS and TERRA. Circular dichroism is used to probe the influence of the cations on the folded fraction and topology of the G4s, while UV thermal melting experiments assess the relative stability of

the structures in each of the tested conditions. The results are compared to the available information on G4 DNAs (*Chapter 2*).

2. Investigate the multimerization tendency of NRAS and TERRA G4 RNAs as a function of cation type and concentration. Polyacrylamide gel electrophoresis, ESI-MS and DLS are the techniques used. The effect of flanking nucleotides is probed by elongating the 18nt-NRAS sequence to 20nt-NRAS, 21nt-NRAS and 22nt-NRAS; as well as cutting down 24nt-TERRA to 21nt-TERRA. The influence of G4 multimerization on *in vitro* experiments is evaluated (*Chapter 3*).

3. Establish a working system to study the G4 NRAS by single-molecule FRET. Observe the folding pathway, kinetics, and dynamics of the RNA in the presence of potassium(I). Assess whether the G4 motif shows any folding intermediates and whether the time scale of dynamics would allow ligand interaction (*Chapter 4*).

The cation interaction studies will allow further insight on the conditions favoring G4 RNA folding and stability, while the multimerization analyses will help in future choices of working sequences for *in vitro* experiments. Finally, with smFRET we aim to understand the mechanisms underlying G4 folding and unfolding, as well as to observe the potential presence of intermediates, and whether the time scale in which G4 motifs are dynamic in physiological conditions would potentially allow ligand interaction *in vivo*.

Chapter 2

Interaction of G-quadruplex RNAs with monovalent and divalent metal ions

2.1 Introduction

Like all nucleic acids, G-quadruplexes (G4s) have a high negative charge due to their phosphate backbone and are therefore sensitive to the presence of metal ions[71] (*Chapter 1*). Moreover, the oxygen lone pairs from the carbonyl groups of G-quartet-forming guanines form a specific cation binding site[70] able to coordinate metal ions[148], which are needed for the folding and stability of the structures. As a consequence of this defined channel binding site, only cations with an adequate size, charge, and dehydration energy are predicted to specifically interact with G4 RNAs[131].

Several studies have reported the interaction between G4 DNA and metal ions[136,211,231-235]; however, less information is available on G4 RNAs[126]. The latter are known to be more stable than their DNA counterparts[128], but no systematic studies with metal ions are available. Therefore, our goal is to explore the formation and stability of two highly-conserved, biologically relevant, RNA quadruplex-forming sequences, 18-nt NRAS and 24-nt TERRA (*Fig. 2.1.*), in the presence of several monovalent and divalent metal ions, namely Li^+, Na^+, K^+, Rb^+, Cs^+, NH_4^+, Mg^{2+}, Ca^{2+}, Sr^{2+}, and Ba^{2+}.

RNA G4s have recently been visualized within the cytoplasm of immobilized human cells by using a G4 structure-specific antibody[114]. This means they are formed *in vivo* in presence of the typical intracellular concentrations of free metal cations, *i.e.*: 5-15 mM Na^+, 140 mM K^+, 0.5 mM Mg^{2+}, 10^{-4} mM Ca^{2+} and $7 \cdot 10^{-5}$ H^+ (pH 7.2)[77]. To address the cation interaction *in vitro*, we use circular dichroism (CD) and thermal difference spectra (TDS), which allow confirmation of the G-quadruplex formation of NRAS and TERRA (*Section 2.2.1.*). Subsequently, G4 folding in each of the ionic conditions is tested, again by CD (*Section 2.2.3.*) and followed by UV thermal melting experiments to assess the relative stability upon each cation addition (*Section 2.2.4.*). Specific attention is directed to Na^+, K^+ and Sr^{2+}, for which titrations are performed (*Sections 2.2.5. and 2.2.6.*).

TERRA	NRAS
5‘-UUA**GGG**UUA**GGG**UUA**GGG**UUA**GGG**-3‘	5‘-**GGG**A**GGG**G**C**G**GG**UCU**GGG**-3‘
	5‘-**GGG**AG**GGG**C**GGG**UCU**GGG**-3‘

Figure 2.1. TERRA and NRAS RNA sequences studied in this work. The guanine bases forming the G-quartets are shown in bold. All other non-guanine bases in between are part of the linking regions, commonly called loops. For the 18-nt NRAS sequence, two folds are possible, depending on which of the guanines from the 4-G run is involved in loop formation.

The effect of different buffers is also discussed (*Section 2.2.2.*), as well as G4 folding in presence of transition metal ions, Fe^{2+}, Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, and Mn^{2+} (*Section 2.2.7.*). ^{1}H NMR spectroscopy is attempted under several conditions (*Section 2.2.8.*) focusing on the imino proton region, which shows characteristic peaks for G4 nucleic acids.

We compare all our findings on the metal ion dependency of the G4 RNA sequences with the existing information on G4 DNAs (*Section 2.2.9.*), giving a detailed view on the metal ion dependency of G-quadruplex formation in RNA.

2.2 Results and discussion

2.2.1 Parallel G4 formation confirmed by CD and TDS

To confirm that the NRAS and TERRA sequences are folding into G4 motifs, circular dichroism and thermal difference spectra are recorded upon K^+ addition. Aware of the stabilizing influence of this cation on G4 DNA[131], and the tendency of G4 RNA to be always parallel-stranded[128], the appearance of characteristic signals for parallel G-quadruplex structures is predicted in both types of spectroscopy.

CD shows a negative minimum at around 242 nm and a positive maximum at 264-265 nm *(Fig. 2.2.a*), typical of parallel G4 formation[236]. Thermal difference spectra also display the characteristic signals for G4 folding: one positive band at 240-243 nm, another positive band at 273-275 nm and one negative band at 295 nm *(Fig. 2.2.b*). This last band corresponds to the stacking between G-quartets[195], confirming that both 18-nt NRAS and 24-nt TERRA fold into G-quadruplex structures.

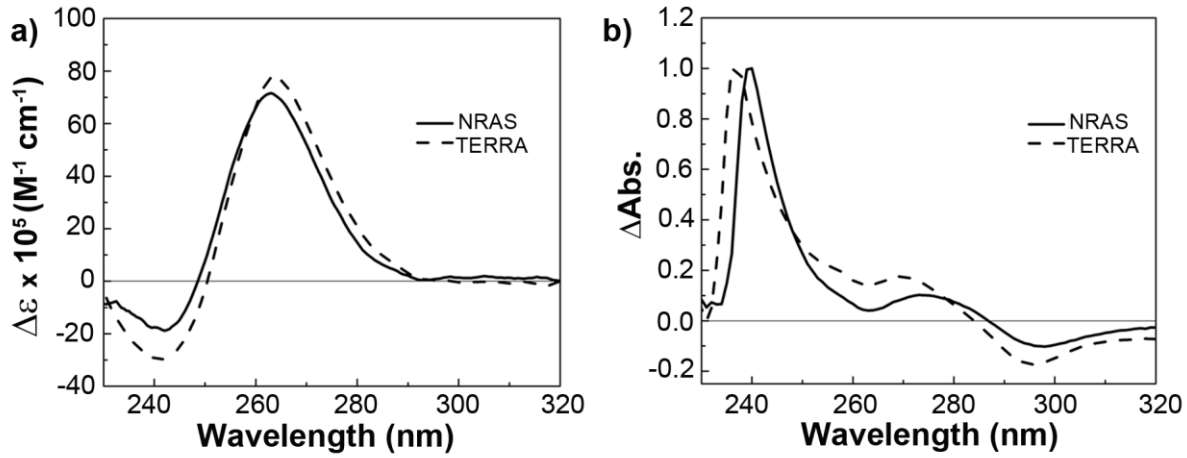


Figure 2.2. NRAS and TERRA RNAs fold into G4 structures. a) CD spectra showing the characteristics of a parallel G4: a negative signal at around 242 nm and a positive signal at around 265 nm. The RNA concentration is 10 µM in 10 mM lithium(I) MOPS pH 7.4 with 20 mM K^+; b) TDS of the NRAS and TERRA RNAs at 10 µM in 10 mM lithium(I) MOPS with 20 mM K^+, showing the characteristic G4 pattern. Both CD and TDS signatures have been normalized and are therefore concentration-independent.

2.2.2 Effect of the buffer conditions on G4 RNA folding and stability

To check whether the RNA is folded without addition of cations, urea (7 M) was used as a denaturing agent to disrupt hydrogen bonding by interacting with the polar, amide-like surfaces of G, C and U bases[237]. In the presence of 10 mM EDTA to chelate any residual mono- and divalent metal ions, CD was measured of both RNAs in urea and compared to water-only conditions. CD intensities are only slightly lower in denaturing agent, for both NRAS and TERRA (*Fig. 2.3.*).

We interpret the small changes between NRAS and TERRA in water and NRAS and TERRA in denaturing medium (urea and EDTA) to be a confirmation that both RNAs were successfully obtained virtually free of cations after desalting, so that their folded fractions were low when dissolved in water. The fact that slight CD signals are observed is probably due to a pre-stacking of the guanine bases, as G-rich sequences are unlikely to result in a completely linear structure, even when unfolded[131]. However, the trapping of traces of monovalent cations in the RNA, even after desalting procedures, cannot be completely excluded. All subsequently observed effects can be interpreted to be due to interactions with the introduced metal ions (*vide infra*). Moreover, TERRA is always more folded than NRAS in water-only conditions (*Fig. 2.3. and 2.5.*).

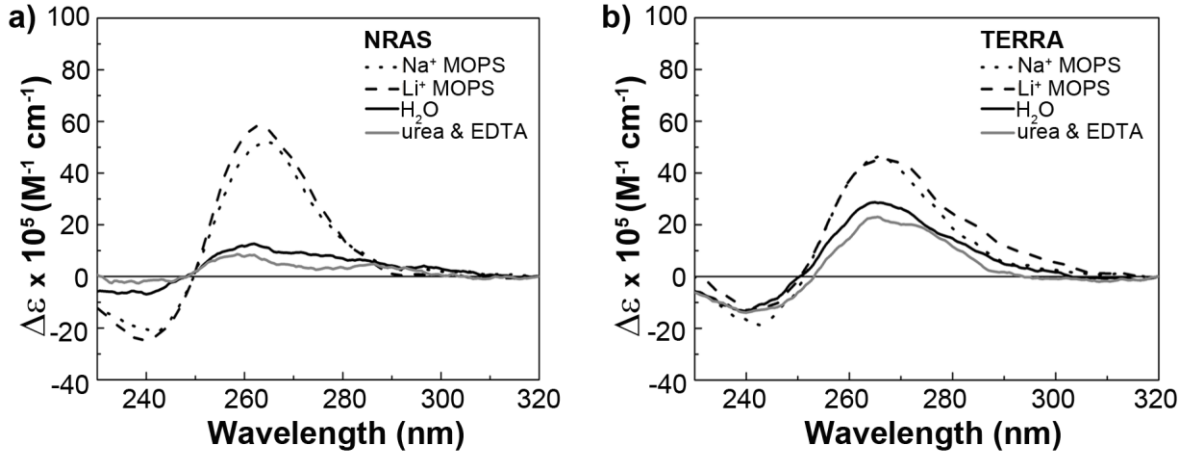


Figure 2.3. CD spectra of G4 RNA in different buffers. Comparison of the G4 RNAs in 1 mM sodium(I) MOPS, 10 mM lithium(I) MOPS (both pH 7.4), and water-only conditions. The RNAs are also measured in 7 M urea with 10 mM EDTA to have a control in which the sequences are expected to be fully denatured. [RNA] = 10 µM, M^+ only from each buffer.

A metal-free MOPS buffer at pH 7.4 containing tetrabutylammonium ions, which are known to be too bulky to fit in between the G-quartets[238], was subsequently tested. This buffer yields non-reproducible and unstable CD signatures in our hands (*Appendix, Fig. A.1.*) and was therefore discarded.

Table 2.1. Melting temperatures in different buffer conditions. Melting temperatures for NRAS and TERRA measured under different buffer conditions.

Buffer (pH 7.4)	M^+ (20 mM)	NRAS T_m (°C)	TERRA T_m (°C)
– (H_2O, pH 6.4)	–	47±1	52±1
– (H_2O, pH 6.4)	Li^+	49±2	51±3
– (H_2O, pH 6.4)	K^+	76±1	65±2
10 mM lithium(I) MOPS	–	50±1	52±1
10 mM lithium(I) MOPS	Li^+	49±1	53±1
10 mM lithium(I) MOPS	K^+	75±1	64±1
1 mM sodium(I) MOPS	–	52±1	47±1
1 mM sodium(I) MOPS	Li^+	50±1	54±1
1 mM sodium(I) MOPS	K^+	76±1	65±1
10 mM lithium(I) cacodylate	–	54±1	53±2
10 mM lithium(I) cacodylate	Li^+	50±3	52±2
10 mM lithium(I) cacodylate	K^+	75±1	66±1

Finally, addition of 1 mM Na^+ or 10 mM Li^+ with the MOPS buffer is sufficient to promote the folding of G4 RNAs (*Fig. 2.3.*). In general, lithium(I) MOPS will be preferred, as Li^+ is not expected to stabilize G4s[131], and its CD signal is lower and more reproducible than that of TBA MOPS.

MOPS is chosen as a buffer instead of phosphate or cacodylate, because of its low coordinating character[239], despite the temperature dependence of its pH[240] (*Appendix, Fig. A.2.*). G4 formation is known to be pH-independent in the 4.5-7.5 range[241]. To confirm that there is no effect from using MOPS, control experiments in a non-temperature-dependent buffer are performed in 10 mM lithium(I) cacodylate (*Table 2.1.*). The obtained T_m values are identical within experimental error for both RNAs in all tested buffers, including water, 10 mM lithium(I) MOPS, 1 mM sodium(I) MOPS and 10 mM lithium(I) cacodylate (MOPS and cacodylate, pH 7.4, and water, measured pH 6.4). The data obtained upon addition of 20 mM of either Li^+ or K^+ in all buffer conditions is also included in *Table 2.1*.

2.2.3 Folding in different metal ion conditions followed by circular dichroism

To follow G4 folding, 20 mM of several monovalent cations (Li^+, Na^+, K^+, Rb^+, Cs^+, NH_4^+) are added to the RNAs, which have been previously dissolved and annealed in water. Nitrate salts are used, instead of the most common chloride analogues, in order to have a non-coordinating

anion that will not interfere with G4 formation[242]. A positive maximum signal at 264-265 nm and a negative minimum signal at 242 nm are seen in all tested conditions (*Fig. 2.4.*). These signals correspond to the characteristic signature for a parallel-stranded G4 RNA[243]. A third peak typical for parallel G4 RNAs (positive, around 210 nm) is not seen in our experiments because of the high nitrate absorption between 200 and 230 nm[244].

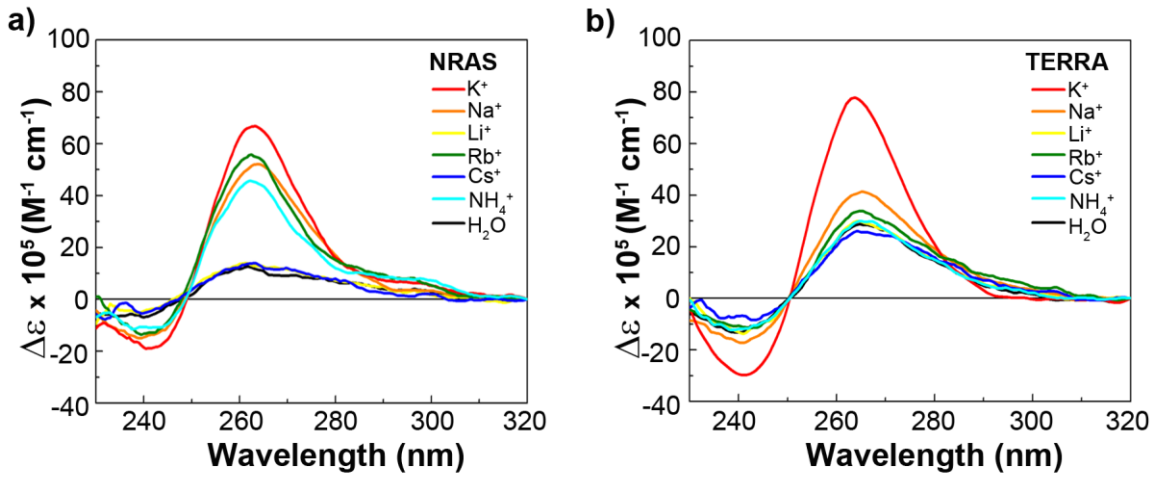


Figure 2.4. CD spectra of G4 RNA in water upon M^+ addition. CD spectra of a) NRAS and b) TERRA RNAs under the following conditions: [RNA] = 10 µM, $[M^+]_{added}$ = 20 mM, in water.

All changes in CD intensities at 265 nm in the different cationic conditions are plotted in *Fig. 2.5.* and expressed as Δ(Δ*ε*), *i.e.* the change in molar CD. With the NRAS sequence the strongest increase in CD signal is observed for K^+ compared to the RNAs in water (*Figs. 2.4.a and 2.5.a*). Na^+, Rb^+, and NH_4^+ yield an increase in the characteristic signals of the parallel CD signature, albeit to a lower extent (*Figs. 2.4. and 2.5.a*). Finally, Li^+ and Cs^+ have no effect at all on the CD signals, their signals superposing with the water-only control.

In the case of TERRA, the most significant Δ(Δ*ε*) is observed in the presence of K^+, the effect of which is considerably higher than that of any other monovalent cation (*Figs. 2.4.b and 2.5.b*). The addition of Na^+ and Rb^+ result in a very moderate increase in the CD signals, while negligible effects are seen upon addition of NH_4^+, Li^+, and Cs^+.2

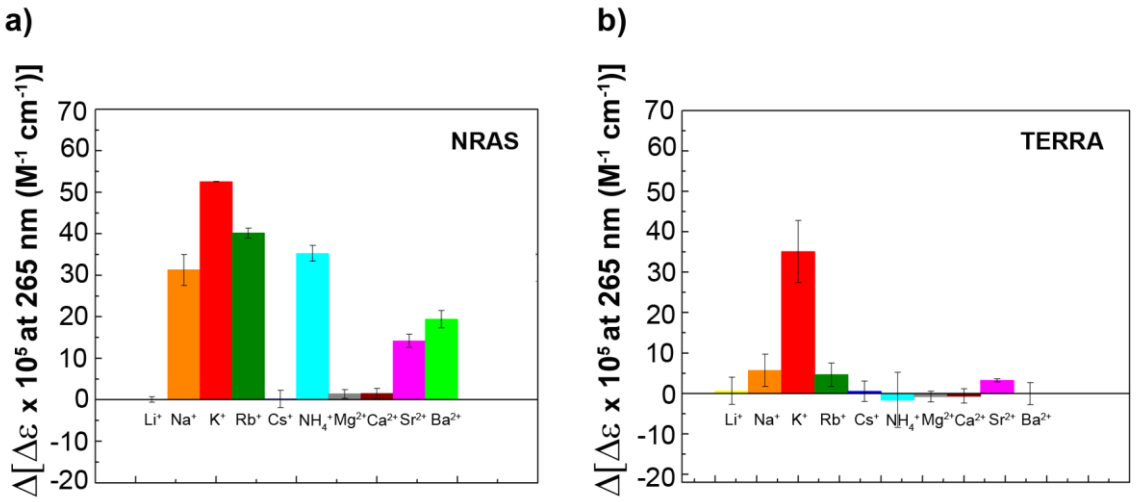


Figure 2.5. Effect of M^{n+} on the folded fraction of G4 RNA. Circular dichroism intensity increment, Δ(Δε), at 265 nm for a) NRAS and b) TERRA G4 RNAs, in relation to the RNAs in water only. [RNA] = 10 μM, $[M^{+}]_{added}$ = 20 mM, $[M^{2+}]_{added}$ = 0.1 mM.

Monovalent cation interaction is further studied in 1 mM sodium(I) MOPS. The addition of 1 mM Na^+ with the buffer is already enough to favor the folding into a G4 compared to water-only conditions, as described above (*Figs. 2.3. and 2.6.*). Further addition of monovalent cations does not, for the most part, result in any further increase of the CD signal at 265 nm. Only K^+ yields a more intense signal, with a comparable Δ(Δε) for both NRAS and TERRA (*Fig. 2.6.*), indicative of a higher folded fraction even under th2e sodium(I) buffer conditions. These observations, in which K^+ is the only monovalent ion tested capable of displacing Na^+ from the MOPS buffer[245], are in agreement with the data from water-only conditions, where K^+ results in the strongest CD increase (*Fig. 2.4. and 2.5.*).

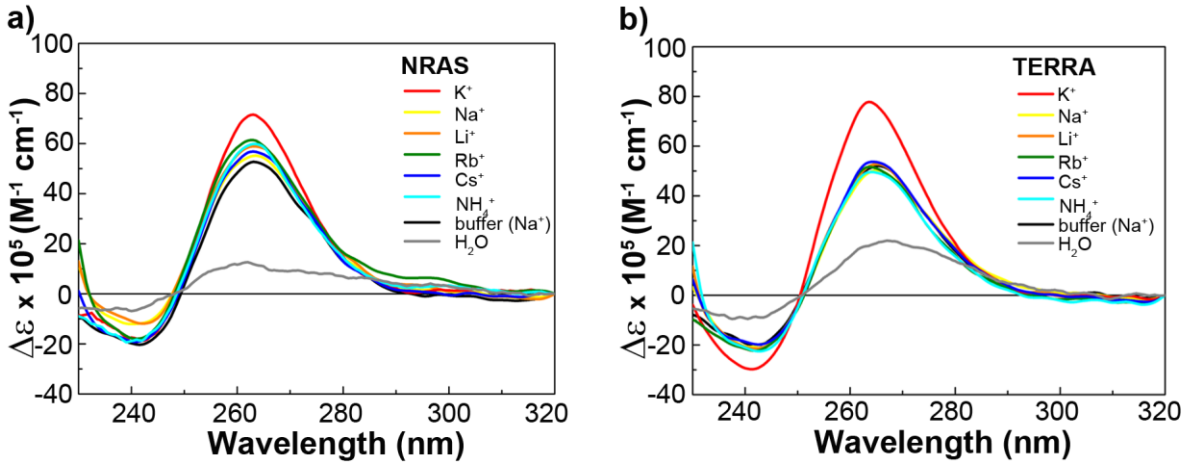


Figure 2.6. CD spectra of G4 RNA in sodium(I) MOPS upon M^+ addition. CD spectra of a) NRAS and b) TERRA RNAs, in the following conditions: [RNA] = 10 μM, $[M^{+}]_{added}$ = 20 mM, in 1 mM sodium(I) MOPS pH 7.4.

The influence of divalent cations is tested in the same manner. Addition of 20 mM M^{2+} (Mg^{2+}, Ca^{2+}, Sr^{2+}, Ba^{2+}) causes the CD signal to decrease strongly (*Fig. 2.7.*). The resulting shape can no longer be unambiguously assigned to a parallel G-quadruplex motif (*Fig. 2.7.*), and a slight shift in the wavelength position of the maxima is observed. Divalent metal ions are known to be able to interact with the guanine functional groups in2volved in G-quartet hydrogen bonding, which might result in G4 unfolding[246,274]. To prevent disturbance of G4 folding CD spectra are measured at 0, 0.1, 1, 2 and 5 mM M^{2+} (*Fig. 2.8.*).

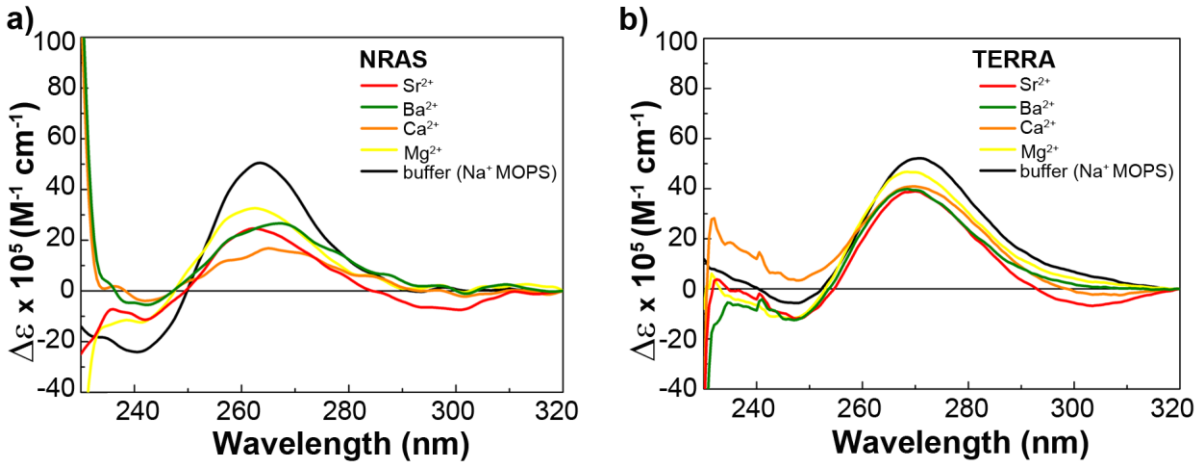


Figure 2.7. CD spectra of G4 RNA in sodium(I) MOPS upon M^{2+} addition. CD spectra of a) NRAS and b) TERRA RNAs, in the following conditions: [RNA] = 10 µM, $[M^{2+}]_{added}$ = 20 mM, buffer = 1 mM sodium(I) MOPS, pH 7.4.

A positive peak at 264-265 nm, and a negative peak at 242 nm consistent with parallel G4 formation[243] are observed under all low [M^{2+}] conditions (*the behavior at 0.1 and 5 mM is shown in Fig. 2.8. and the Δ(Δε) at 0.1 mM are plotted in Fig. 2.5.*). The intensity of the NRAS G4 signals increases upon addition of 0.1 mM Sr^{2+} or Ba^{2+} (*Fig. 2.5. and 2.8.a*), while 5 mM are needed to see the response to Ca^{2+} and Mg^{2+}. At both concentrations, the Δ(Δε) is greatest for Ba^{2+}, closely followed by Sr^{2+}.

At 0.1 mM, only Sr^{2+} exerts a noticeable effect on TERRA by increasing the intensity of the signal at 265 nm (*Fig. 2.5. and 2.8.b*). Upon addition of 5 mM M^{2+} an increase of the CD intensity is observed for all cations except for Mg^{2+}, which yields the same spectrum as the RNA in water-only.

A higher intensity of the signals in the CD spectra indicates that a greater fraction of the RNA is folded into a G4 structure. Therefore, the observed Δ(Δε) can be interpreted in terms of the folding to G-quadruplex motifs being promoted to a higher or lower extent in the different mono- and divalent cations tested. For NRAS, the folding to a G4 structure is highly boosted upon addition of K^+, Sr^{2+}, or Ba^{2+}, favored in the presence of Na^+, Rb^+, NH_4^+, Ca^{2+}, and Mg^{2+}, and no further encouraged by Li^+ or Cs^+. In contrast, the G4 folding of TERRA is most

increased by K^+ and Sr^{2+}, slightly promoted by Na^+, Rb^+, Ba^{2+}, and Ca^{2+}, and unaffected by Li^+, Cs^+, NH_4^+, and Mg^{2+}.

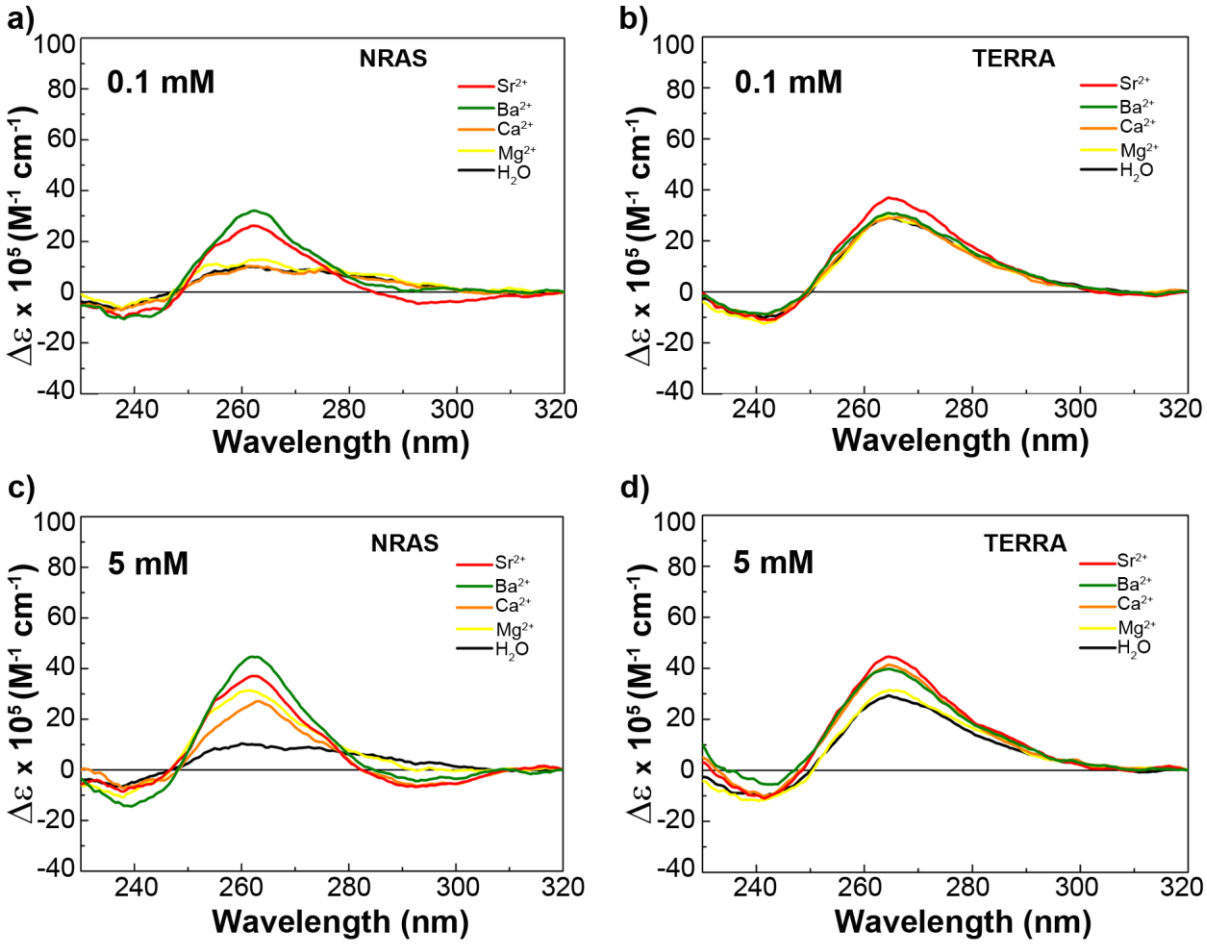


Figure 2.8. CD spectra of G4 RNA in water upon M^{2+} addition. CD spectra of a) NRAS with 0.1 mM M^{2+}, b) TERRA with 0.1 mM M^{2+}, c) NRAS with 5 mM M^{2+}, d) TERRA with 5 mM M^{2+}. [RNA] = 10 µM, $[M^{2+}]_{added}$ = 0.1 mM or 5 mM, in water.

K^+, Na^+, and Sr^{2+} are already known to promote the folding of G4 DNAs[128], as they do for both G4 RNAs studied herein. The fact that no effect upon addition of Li^+ or Cs^+ is observed with any of the sequences is also in agreement with literature[231,247] and with the fact that these cations do not possess a compatible size with the G4 binding site (*Table 2.2.*). Some differences in cation dependence are observed between NRAS and TERRA. Indeed, NRAS is more sensitive to both monovalent and divalent cation effects (*Fig. 2.4. and Tables 2.3. and 2.4.*), presumably due to its shorter loops (*Fig. 2.1.*) creating a higher repulsion in the RNA backbone. The folding of NRAS is also promoted by Ba^{2+} to a higher extent than the folding of TERRA. This stabilization by Ba^{2+} is already known for the thrombin aptamer, an intramolecular G4 DNA[248], and has been described within the crystal structures of other G4 DNAs[249,250].

Table 2.2. Relevant characteristics of the tested monovalent and divalent cations. Ionic radii and dehydration energies relevant in terms of the cation interaction with G4 RNA. The distance between two G-quartet planes in a parallel G4 is 3.13 Å[70]. The cations having the most ideal characteristics for G4 interaction are marked in gray.

Metal ion	Ionic radius (Å)[a]	Hydrated radius (Å)[b]	$\Delta H°hyd/z^2$ (KJ/mol)[c]	Bound H_2O[d]
Li^+	0.90	3.40	−519	22
Na^+	1.16	2.76	−409	13
K^+	1.33	2.32	−322	7
Rb^+	1.66	2.28	−293	7
Cs^+	1.81	2.28	−264	6
NH_4^+	1.43	5.35	−307	5
Mg^{2+}	0.65	7.00	−480	6 / 36[e]
Ca^{2+}	0.99	4.12	−394	6 / 29
Sr^{2+}	1.18 (6L[f]) 1.26 (8L[f])	4.12	−361	6 / 29
Ba^{2+}	1.35	5.90	−326	8 / 28

[a] Refs.[70,124,251,252]; [b] refs.[253-257]; [c] experimental hydration enthalpies, charge normalized[258,259]; [d] bound water molecules[260-263]; [e] number of water molecules in the first coordination sphere / total number of water molecules; [f] Sr^{2+} can bind either 6 or 8 ligands (6L and 8L), with its ionic radius differing depending on its coordination.

2.2.4 Relative G4 RNA stability measured by UV thermal melting

UV thermal melting experiments are performed to assess the relative stability under different cation conditions. The melting temperatures (T_m) calculated from the recorded melting and annealing curves are a reliable measure of the thermal stability for comparative purposes[239].

Preliminary experiments in 20 mM K^+ show a full melting curve, with an observable melting transition. This concentration is therefore selected as the cation concentration for the comparison of melting curves with other metal ions. Addition of most of the tested monovalent cations does not increase the melting temperature compared to the RNAs in water-only conditions nor in non-stabilizing buffer (*Table 2.3. and Fig. 2.9.*). The buffer contains 10 mM lithium(I), as Li^+ has been reported not to stabilize G4 formation[131] (*Table 2.1.*). Measurements with the corresponding DNA sequences, dNRAS and htelo, are undertaken for comparison.

The two RNA sequences, NRAS and TERRA, are already moderately stable in water only, with T_m's of 45±3 and 51±1 °C, respectively (*Table 2.3.*). This is in agreement with the fact that water molecules are known to have some stabilizing effect on G4s[264]. Such a high melting temperature without the presence of monovalent cations has already been observed for a model G4 RNA motif, intramolecular and containing 3 G-quartets: 5'-(GGGU)$_3$GGG-3', with

T_m = 50 °C[128]. Moreover, molecular dynamics simulations have shown that G4s can be stable without coordinated cations. In that case, water molecules occupy the empty coordination sites and G-quartet guanosines rotate to move their O6 apart from each other, thus minimizing electronic repulsions[265]. However, as mentioned above, residual traces of monovalent cations in the desalted RNA samples cannot be completely excluded.

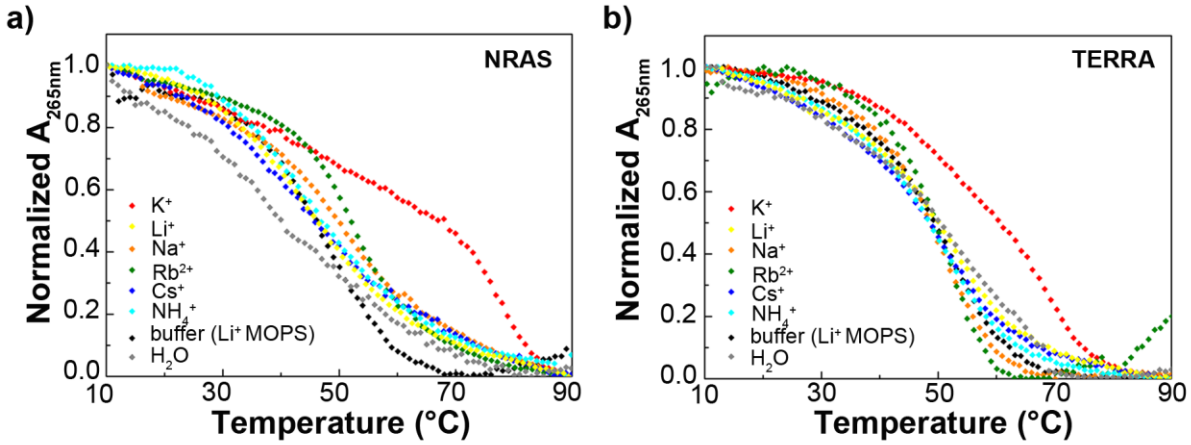


Figure 2.9. G4 RNA melting curves in 20 mM M^+. Example of UV melting curves in the presence of Li^+, Na^+, K^+, Rb^+, Cs^+, and NH_4^+. [RNA] = 4 µM, $[M^+]_{added}$ = 20 mM, buffer = 10 mM lithium(I) MOPS, pH 7.4.

The addition of 10 mM Li^+ with the lithium(I) MOPS buffer does not affect the recorded melting temperatures. Li^+ has been reported not to stabilize G4[131] because of its small ionic radius and high dehydration energy (*Table 2.2.*)[225,266]. Interestingly, the addition of 10 mM lithium(I) MOPS, albeit without affecting the stability of the G4s, does increase the folded fraction of RNA as seen in the CD spectra in *Fig. 2.3.*, possibly by a simple charge-balancing effect.

Upon addition of monovalent cations, both NRAS and TERRA are only stabilized by K^+ (*Tables 2.3. and 2.4.*) and an increase of $\Delta T_{m(TERRA)}$ = 12±1 °C and $\Delta T_{m(NRAS)}$ = 25±1 °C is observed for the measurements performed in lithium(I) MOPS buffer with 20 mM K^+. This corresponds well to the reported observations for G4 DNA, in which K^+ is described to have a high selectivity over other monovalent cations[267]. The obtained $\Delta T_{m(TERRA)}$ is in good agreement with the reported $\Delta T_{m(TERRA),lit}$ = 12.6 °C[268]. TERRA G4 melting in water-only and in K^+ is also followed by CD (*Appendix, Figs. A.3.-A.5.*) yielding comparable results to the UV measurements.

All other M^+ result in T_m's that are comparable to the temperatures in water / buffer only (*Table 2.3.*). In presence of K^+, the NRAS G4 is more stable than the TERRA G4 (*vide supra and Table 2.3.*). This can be related to the shorter loops of the NRAS sequence (1-2-3 nt or 2-1-3 nt) compared to TERRA (3-3-3 nt) (*Fig. 2.1.*), as it is known that thermal stability decreases with an increase in the total loop length[120].

No melting transitions are observed for both sequences in the presence of 20 mM divalent cations (only for TERRA in Ba^{2+}; *Fig. 2.10.*). This is in agreement with CD measurements that show no clear G4 signals at the same divalent cation concentration (*vide supra and Fig. 2.7.*). The melting experiments are subsequently performed in the presence of 0.1 mM of M^{2+} (*Fig. 2.11.*), as no analyzable melting curves are observed at 5 mM. This absence of melting at higher divalent cation concentrations might be due to the fact that the G4s are too stable to be unfolded at 90 °C. The presence of multimers is also suspected (Chapter 3).

Upon addition of 0.1 mM of M^{2+}, Sr^{2+} gives the strongest stabilization effect as compared to the RNAs in buffer-only conditions with a similar increase in melting temperature for both TERRA and NRAS, $\Delta T_{m(TERRA)}$ = 22±1 °C and $\Delta T_{m(NRAS)}$ = 25±1 °C (*Table 2.4. and Fig. 2.11.*).

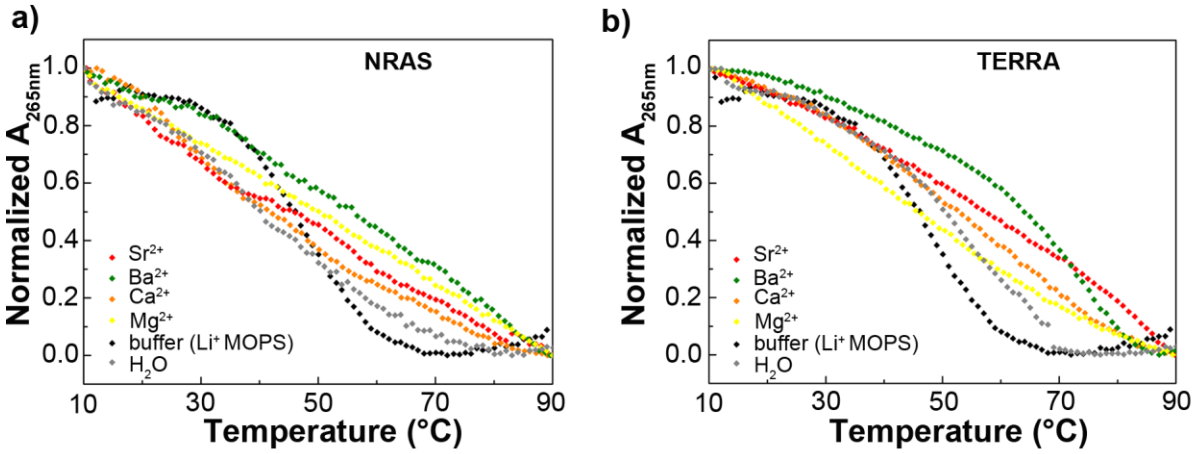


Figure 2.10. G4 RNA melting curves in 20 mM M^{2+}. Example of UV melting curves in the presence of Sr^{2+}, Ba^{2+}, Ca^{2+}, and Mg^{2+}. [RNA] = 4 µM, $[M^{2+}]_{added}$ = 20 mM, buffer = 10 mM lithium(I) MOPS, pH 7.4.

For NRAS, a pattern of stability in divalent cations is observed as Sr^{2+} >> Ba^{2+} > Ca^{2+} > Mg^{2+} (*Table 2.4.*), following the sizes and dehydration energies of the metal ions (*Table 2.2.*). In contrast, TERRA yields a similar behavior for Ba^{2+}, Ca^{2+}, and Mg^{2+}, with a significant stabilization only in Sr^{2+}, which in fact has a stronger effect than K^+, presumably due to the higher affinity of a divalent cation for RNA *versus* a monovalent cation.

Mg^{2+} has been reported to have a destabilizing effect on telomeric G4 DNA[70] while the effect on promoter G4 DNA is stabilizing[264]. For the G4 RNAs studied herein no effect is observed as the recorded melting temperature is of the same order as in the presence of 10 mM Li^+ from the buffer. This lack of Mg^{2+} effect is possibly due to a higher stability of the RNA G-quadruplexes in water-only conditions (*vide supra and Table 2.3.*).

Table 2.3. G4 RNA stability in M^+. Melting temperatures for NRAS and TERRA under different monovalent cation conditions in 10 mM lithium(I) MOPS pH 7.4. The DNA equivalents (dNRAS and htelo), either measured in this work or reported from the literature, are included for comparison. [RNA] = 4 µM, $[M^+]_{added}$ = 20 mM. The most stable condition is marked in gray.

M^+ (20 mM)	NRAS T_m (°C)	dNRAS T_m (°C)	TERRA T_m (°C)	htelo T_m (°C)	htelo – lit.[a] T_m (°C)
H_2O only	47±1	27±1	52±1	24±1	–
10 mM Li^+ (from buffer)	50±1	29±3	52±1	24±1	–
Li^+	49±1	26±1	53±1	25±2	23.0±0.2 (10 mM) 27.8±0.2 (49 mM)
Na^+	50±1	n/a	53±1	n/a	42.4±0.2 (49 mM)
K^+	75±1 68±1[b]	60±1	64±1 58±1[b]	51±1	44.3±0.5 (10 mM) 50.2±0.2 (49 mM)
Rb^+	53±1	n/a	52±1	n/a	26.4±0.3 (10 mM) 34.8±0.1 (49 mM)
Cs^+	47±1	n/a	51±1	n/a	27.4±0.3 (180 mM)
NH_4^+	47±2	n/a	51±1	n/a	–

[a] Measured with the telomeric DNA sequence $(TTAGGG)_4$[223]; [b] Measured in 5 mM K^+ [120].

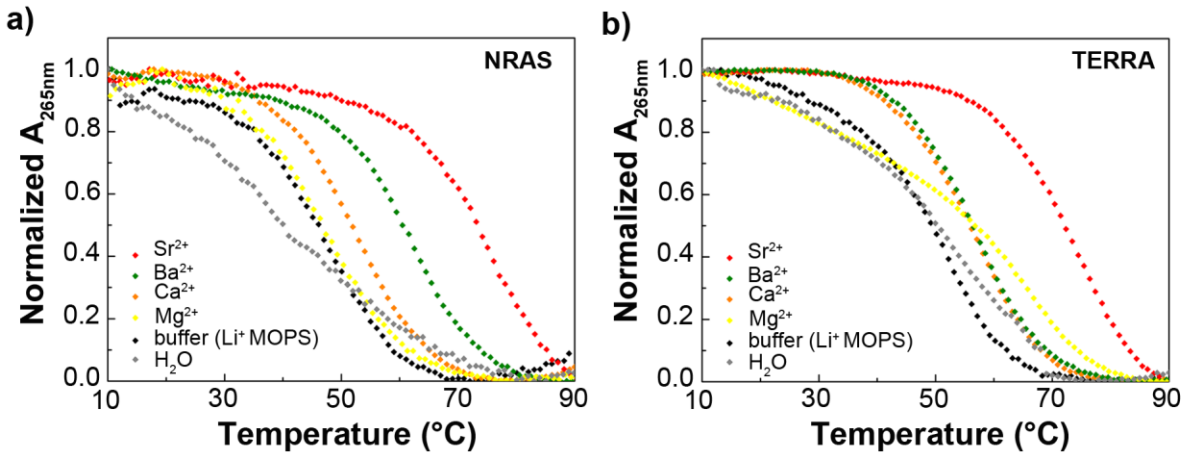


Figure 2.11. G4 RNA melting curves in 0.1 mM M^{2+}. Example of UV melting curves in the presence of Sr^{2+}, Ba^{2+}, Ca^{2+}, and Mg^{2+}. [RNA] = 4 µM, $[M^{2+}]_{added}$ = 0.1 mM, buffer = 10 mM lithium(I) MOPS, pH 7.4.

Ba^{2+} appears to have a stronger effect on NRAS than on TERRA (*Table 2.4. and Fig. 2.8.*) yielding a more stable G4 with a non-negligible $\Delta T_{m(NRAS)}$ = 6±1 °C compared to the buffer conditions while having no significant effect on TERRA. This observation can presumably be related to the intrinsically higher stability of NRAS in presence of cations due to the shorter loops of the structure (*vide supra and Fig. 2.1.*), which create a higher repulsion in the RNA backbone.

To test for the influence of RNA concentration on the stability of the G4 RNAs, melting transitions are recorded from 1 to 50 µM RNA concentration under constant 20 mM K^+ or 0.1 mM Sr^{2+}. The obtained melting temperatures are constant within the error limits in the presence of K^+ at low RNA concentration (up to 10 µM), while slightly increasing at higher RNA concentration (*Fig. 2.12.a*). In the presence of Sr^{2+}, T_m's are the same within the error limits for TERRA. For NRAS, they start increasing at 4 µM (*Fig. 2.12.b*), which suggests that this G4 RNA might be multimerizing / aggregating (*Chapter 3*).

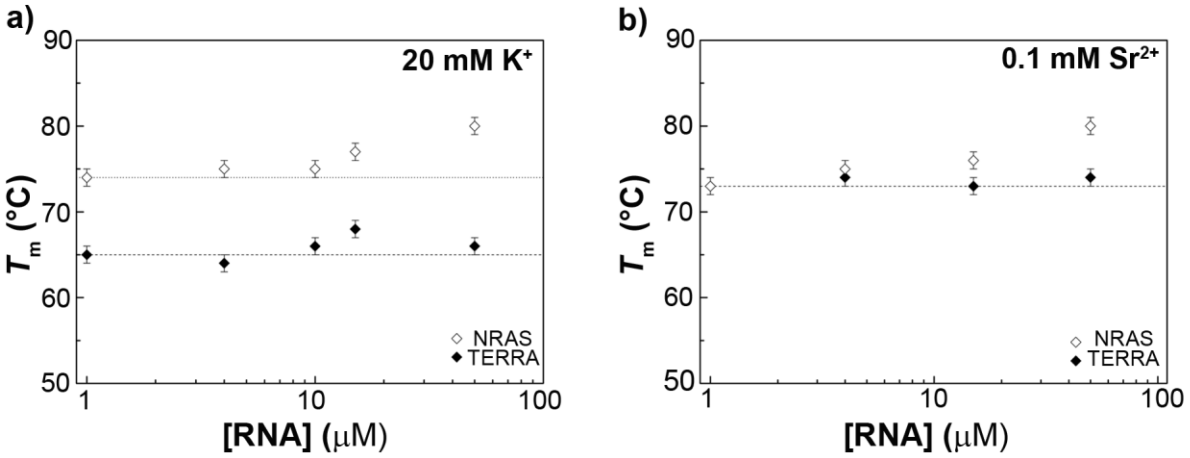


Figure 2.12. The RNAs fold into intramolecular G4s. Melting temperatures in 10 mM lithium(I) MOPS buffer, pH 7.4, with a) 20 mM K^+ or b) 0.1 mM Sr^{2+}, as a function of RNA concentration (1-50 µM). The lines are drawn at the coordinate of the first RNA concentration for visualization.

The fact that the melting temperatures are independent of RNA concentration, which is the case in both K^+ and Sr^{2+} at low RNA concentrations, indicates that G4 folding is intramolecular. At higher RNA concentrations the T_m's start to deviate, probably due to the formation of G4 multimers. The deviation is more pronounced for NRAS, which corresponds with the expected higher tendency towards aggregation in absence of flanking nucleotides[117]. A detailed view on the G4 RNAs aggregation behavior is presented in *Chapter 3*.

Table 2.4. G4 stability in M^{2+}. Melting temperatures for NRAS and TERRA under different monovalent cation conditions in 10 mM lithium(I) MOPS pH 7.4. The DNA equivalents (dNRAS and htelo), either measured in this work or reported from the literature, are included for comparison. [RNA] = 4 µM, and $[M^{2+}]_{added}$ = 0.1 mM. The most stable condition is marked in gray.

M^{2+} (0.1 mM)	NRAS T_m (°C)	dNRAS T_m (°C)	TERRA T_m (°C)	htelo T_m (°C)	htelo – lit. T_m (°C)
H_2O only	47±1	27±1	52±1	24±1	–
10 mM Li^+ (from buffer)	50±1	29±3	52±1	24±1	–
Mg^{2+}	48±1	n/a	53±1	n/a	–
Ca^{2+}	54±1	n/a	55±1	n/a	–
Sr^{2+}	75±1	55±1	74±1	47±1	73.2±0.3 [a]
Ba^{2+}	56±1	n/a	54±1	n/a	–

[a] Measured in 10 mM Sr^{2+} with the telomeric G4 DNA sequence $(TTAGGG)_4$[223]. This temperature is higher than the one measured by us, as our experiments were performed at 0.1 mM Sr^{2+}.

The melting temperatures give a measure of the relative stability of the G4 RNAs. Therefore, we can hypothesize that these motifs would be thermodynamically stable if formed *in vivo* at 37 °C, as all calculated T_m's are higher than this temperature. K^+ and Sr^{2+} are the only cations that result in a significant increase in melting temperatures. They both have ideal ionic radii for the cavity between two quartets in the G4 structure, as well as small dehydration energies (*Table 2.2.*). This allows them to eliminate their surrounding hydration waters and to coordinate to eight O6 oxygens from the G-quartet guanines fitting within two G-quartets[269] in a bipyramidal antiprismatic geometry[251]. This direct coordination of specifically bound cations[247] is what confers a high stabilization to the G4 RNA structures as shown by the melting temperatures reported herein (*Tables 2.3. and 2.4.*). Other metal ions either do not have an adequate size to fit into the binding cavity (e.g. Li^+ or Cs^+) or else are harder to dehydrate (e.g. Li^+ or Na^+) (*Table 2.2.*).

Finally, any differences in behavior between NRAS and TERRA are not expected to arise from the direct coordination in the internal ion channel, as the core of three G-quartets is likely analogous, but rather from the presence or absence of flanking nucleotides and from the composition and length of the loops (*Fig. 2.1.*), which are also known to be able to electrostatically attract metal ions[264].

Significantly, K^+ is the monovalent cation with the highest intracellular concentration, ca. 140 mM[77], which would indicate the relevance of these motifs *in vivo*. Sr^{2+} has no known biological function and is toxic if inhaled[270]. Its high affinity for G4 nucleic acids has been exploited towards the design of sensitive sensors for its detection[270].

2.2.5 Sodium(I) vs. potassium(I) titrations

To further study the different influence of Na^+ and K^+ on G4 RNA formation and stability, we measure CD spectra of NRAS and TERRA in water and after adding 0.1-150 mM Na^+ or K^+. K^+, as we have shown (*vide supra*), is the most stabilizing M^+ for G4 RNAs while Na^+ has been described to stabilize G4 DNAs[128,231].

The characteristic positive CD signal at 265 nm for a parallel G4 fold shows an increase in intensity upon Na^+ or K^+ addition for both sequences, indicating a shift in equilibrium towards G4 formation (*Fig. 2.13.a*). A moderate addition of Na^+ or K^+ (up to 10 mM of added salt, corresponding already to an excess of cations vs. the RNA) leads to a strong increase of the CD intensity and yields the characteristic signature of a parallel G4 for both RNA sequences. It seems that most of the folding takes place at low salt concentrations (5-10 mM) and is followed by a saturation up to 150 mM (*Fig. 2.13.a*), which indicates that no further folding of the structures is occurring after 10 mM of added M^+.

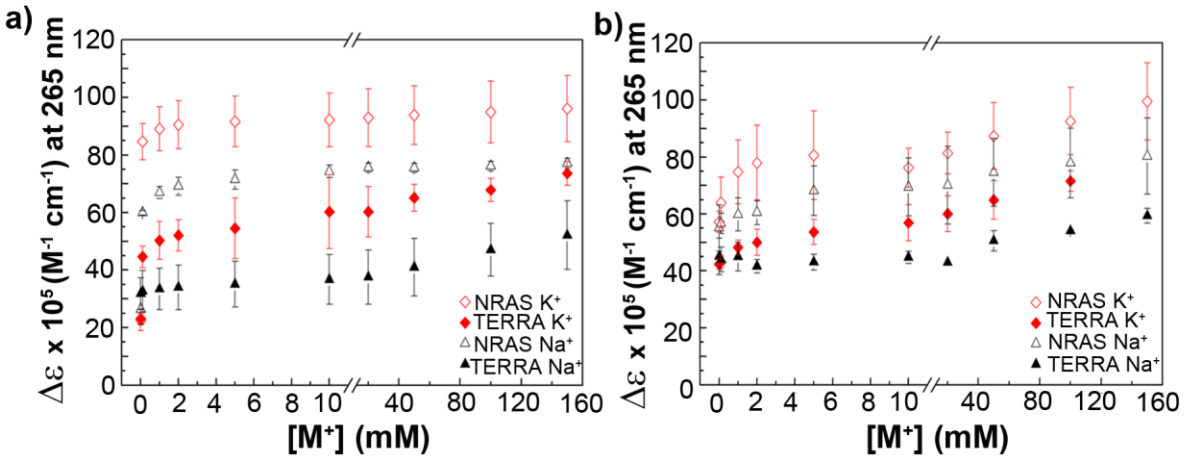


Figure 2.13. G4 RNA titration curves upon Na^+ and K^+ addition. Recorded with circular dichroism for TERRA and NRAS G4 RNAs in a) water or b) 10 mM lithium(I) buffer pH 7.4. A break was introduced in the x-axis scale, in order to better visualize the intensity increase below 10 mM of added cations. [RNA] = 15 µM, $[M^+]_{added}$ = 0.1-150 mM.

The same CD titrations are performed in 10 mM lithium(I) MOPS (*Fig. 2.13.b*). The presence of the monovalent cations from the buffer (10 mM Li^+) is enough to favor the partial folding into a G4 (*Fig. 2.3.*) without K^+ or Na^+ addition, and therefore a more gradual increment in $\Delta(\Delta\varepsilon)$ is observed overall (*Fig. 2.13.b*).

The observed increase in CD signal upon addition of K^+ to RNAs in non-cationic conditions is already known for TERRA and other G4 RNAs[128]. Na^+ addition yields smaller $\Delta(\Delta\varepsilon)$ than K^+ for both RNAs, which indicates that sodium(I) has a smaller effect than potassium(I) on promoting G4 RNA folding.

UV melting experiments are performed with increasing concentrations of Na^+ and K^+ in non-stabilizing 10 mM lithium(I) MOPS buffer to test the cation-dependent thermodynamic stability of the G4s. In K^+, T_m's of both NRAS and TERRA show a significant increase if K^+ is added up to 25 – 30 mM (*Fig. 2.14.a*). Further addition of K^+ until 80 mM causes no further stabilization of the G4s, indicated by constant T_m's (*Fig. 2.14.a*). Upon addition of 100 mM K^+, the RNA structures are too stable to be melted. The strong increase in stability at low K^+ concentrations and the appearance of a plateau after ~ 30 mM K^+ fit to a single exponential model (*Fig. 2.14.b*). This shows the linear relation between log(K^+ concentration) and temperature and indicate that G4 stability is proportional to the K^+ concentration present in solution. Contrarily, the addition of 1-100 mM Na^+ has no effect on the stability of neither NRAS nor TERRA, and the T_m's are constant within error across the concentration range tested (*Fig. 2.14.a*).

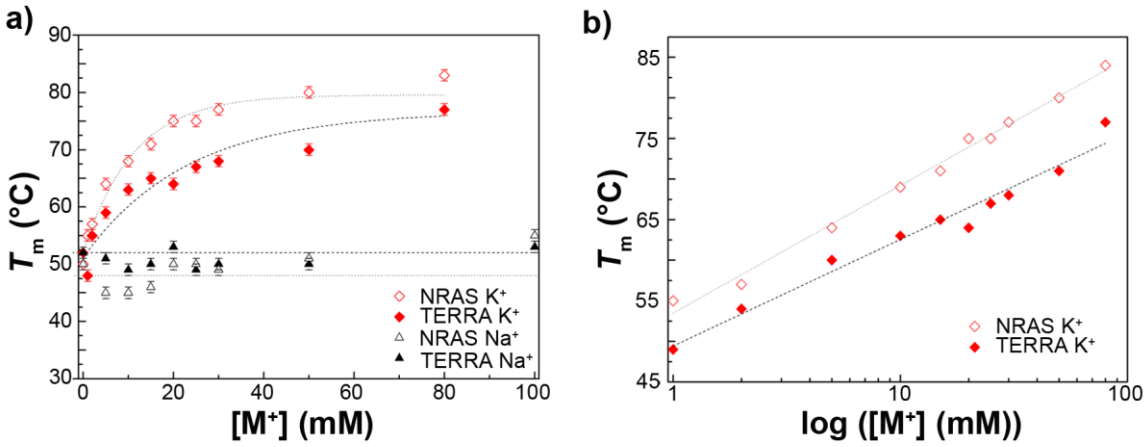


Figure 2.14. G4 RNA stability upon Na^+ and K^+ titrations. a) Titration curves created by plotting the results of UV thermal melting experiments in 1-100 mM Na^+ or 1-80 mM K^+ for TERRA or NRAS G4 RNAs. [RNA] = 4 µM, buffer = 10 mM lithium(I) MOPS, pH 7.4. b) Direct correlation of G4 stability with K^+ concentration.

All in all, K^+ has a strong stabilizing effect on the G4 structures, as already hinted by the single T_m measurements at 20 mM (*Table 2.3.*). On the other hand, Na^+ does not have any stabilizing effect on the G4 RNA motifs, as there is no further increase in T_m with respect to 10 mM lithium(I) MOPS buffer (*Fig. 2.14.a*) or to water-only conditions (*Table 2.3.*). The insensitivity towards Na^+ might indicate that this cation is acting only as a non-specific ion binder without direct coordination to the G-quartets and thus does not lead to any conformational changes in respect to the RNAs in buffer / water-only conditions[247].

We can therefore conclude that both Na^+ and K^+ increase the fraction of folded G4 upon their addition (the more so in the case of K^+) but only K^+ leads at the same time to a stabilization of the quadruplex structures, as observed through a direct correlation between K^+ concentration and T_m increase (*Fig. 2.14.b*). For both NRAS and TERRA, 5-10 mM of M^+ are required for

most of the folding to occur for 15 µM RNA. This represents already a 1:1000-1:2000 cation excess, indicating that G4 folding requires more K^+ than what would be necessary to occupy the channel positions[85]. Indeed, K^+ is known to be able to localize in loops and grooves of G4 motifs, with extensive intramolecular contacts[264].

2.2.6 Strontium(II) titrations

The behavior of Sr^{2+} in promoting G4 folding is further followed by CD titrations owing to its highest stabilizing effect among the tested divalent cations (*vide supra*). Both RNA sequences dissolved in metal-free TBA buffer show the characteristic CD shape of a parallel-stranded G4. Subsequently, upon addition of Sr^{2+}, the intensity of both G4 signals decreases gradually, in direct correlation with the concentration of the divalent cation (*Fig. 2.15.*). This is in agreement with the recorded CD spectra in 20 mM M^{2+} (*vide supra and Fig. 2.7.*).

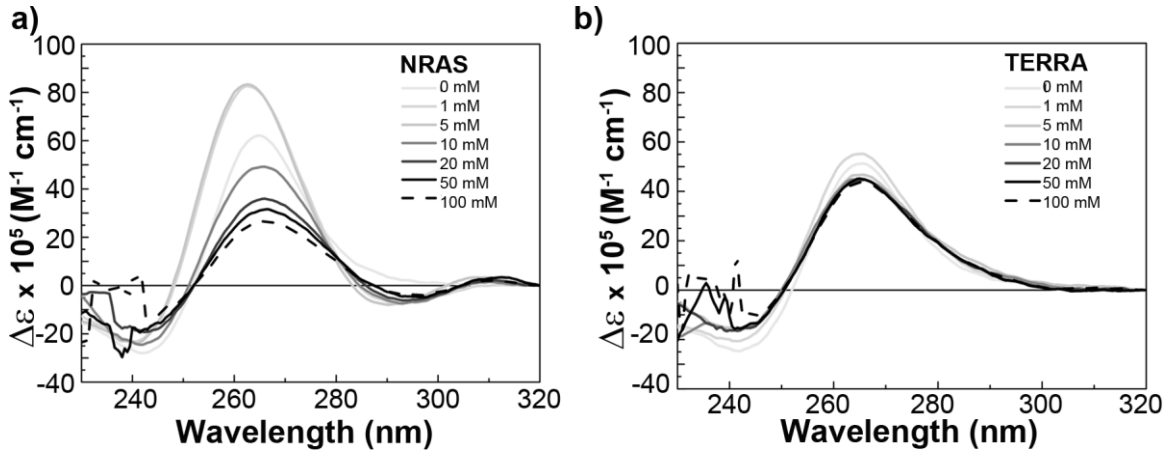


Figure 2.15. G4 RNA CD titrations with Sr^{2+}. a) NRAS and b) TERRA upon titration with Sr^{2+}. [RNA] = 10 µM, $[M^{2+}]_{added}$ = 0-100 mM, in 10 mM TBA MOPS, pH 7.4.

In order to verify that Sr^{2+} is not degrading the RNAs, a 15 µM NRAS sample with 75 mM Sr^{2+} is analyzed by MALDI-MS. The resulting spectrum (*Appendix, Fig. A.17.*) shows a molecular peak at 6220.1 g/mol, corresponding to the 5'-triphosphate full-length RNA (calcd. 6217.5 g/mol), as well as Sr^{2+} adducts containing one, two or four cations. This shows that the decrease in CD signals is not caused by RNA degradation. At the same time, multimerization of G4 monomers in presence of M^{2+} is suspected and will be further explored in *Chapter 3*.

2.2.7 The effect of transition metal ions on G4 RNA

Little is known about the ability of divalent metal ions other than the earth alkali cations to stabilize G4 DNA or RNA. Millimolar concentrations of the transition metal ions Mn^{2+} (4-10 mM), Co^{2+} (0.3-2mM), and Ni^{2+} (0.3-0.8 mM) were previously reported, using CD, to counteract the potassium-induced G4 formation in DNA[271]. The authors reported a shift towards the triplex form in presence of these cations, when a G*G·C-based triple helical

structure in equilibrium with a potassium-induced G4 DNA was studied. Moreover, an excess of Cu^{2+} has been reported to unfold a G4 DNA[272].

Engelhard *et al.* reported the reversible stabilization of G4 DNA binding Cu^{2+} or Ni^{2+}. However, the guanine-rich sequences were modified by adding a pyridyl functionality to the 5' end[273]. The resulting tetramolecular G4 motif was thus able to coordinate the transition metal ion in a square-planar fashion through the monodentate pyridine ligands, forming a "metal-base tetrad". Upon Ni^{2+} or Cu^{2+} binding, the structure was stabilized towards thermal melting, yielding melting temperature increases of ΔT_m = 15 °C and ΔT_m = 20 °C, respectively. The transition metal coordination was reversible upon addition of EDTA as a chelating agent. The authors suggested that this type of metal-carrying G4s might be developed as spectroscopically traceable probes.

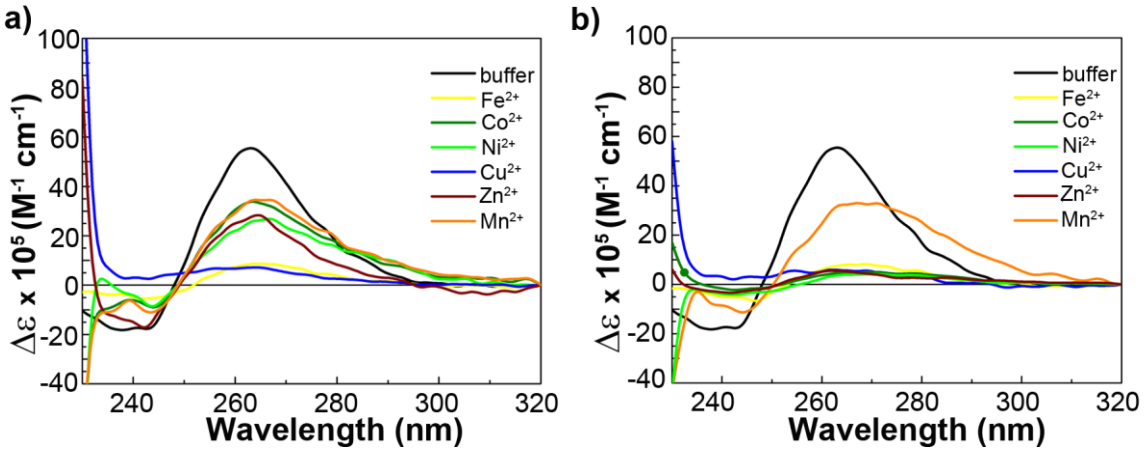


Figure 2.16. CD spectra of G4 RNA with transition metal ions. a) NRAS and b) TERRA RNAs. [RNA] = 10 µM, $[M^{2+}]_{added}$ = 20 mM, in 1 mM sodium(I) MOPS, pH 7.4. Fe^{2+} was measured in 1 mM sodium(I) MOPS pH 6.4 to minimize oxidation.

Herein, the NRAS and TERRA G4 RNAs are followed upon addition of 20 mM of the following transition metal ions: Fe^{2+} (as $Fe(ClO_4)_2$), Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, or Mn^{2+} (as $M^{2+}(NO_3)_2$). The recorded CD spectra are compared to the corresponding RNA sequence dissolved in 10 mM sodium(I) MOPS without further added cations. Fe^{2+} precipitates from the solution as soon as the buffer is added, presumably because the increase in pH, from ~ 6.4 to 7.4, results in Fe^{2+} to Fe^{3+} oxidation and formation of the insoluble $Fe(OH)_3$. Sodium(I) MOPS buffer is subsequently prepared at pH 6.4 to minimize the oxidation of Fe^{2+}. At this lower pH, the solution still appears slightly brownish and no G4 signals are observed. As for Cu^{2+}, it clearly destabilizes G4 folding for both RNAs yielding almost no CD signal (*Fig. 2.16.*).

In comparison to buffer-only conditions, the spectra of NRAS with Co^{2+}, Ni^{2+}, Zn^{2+}, or Mn^{2+}, and of TERRA with Mn^{2+} yield lower CD intensities. The signals are slightly shifted to higher wavelengths, presumably corresponding to a destabilization of the G4 structures[236,274]. TERRA in the presence of Co^{2+}, Ni^{2+}, or Zn^{2+} has completely lost the G4 signals. Due to

precipitation of the transition metal salts at high temperatures, no melting studies are performed.

2.2.8 ^{1}H NMR showing the imino proton region of G4s under different temperature and cation conditions

G-quadruplex formation yields characteristic guanine imino protons (H1), which can be observed by ^{1}H NMR spectroscopy at 9-12 ppm[208] (*Chapter 1 and Fig. 1.19.*). In these structures, the number of observed imino proton resonances correlates with the number of guanine bases involved in G-quartet formation. With the aim of further studying the effect of different cation conditions on the G4 RNA structures, ^{1}H NMR spectra are measured at several K^+ and Sr^{2+} concentrations. The influence of the temperature is also followed, as well as the addition of crowding agents and of a G4 ligand.

Dimeric 12-nt TERRA used as a control

The already reported 12-nt TERRA sequence 5'-UAGGGUUAGGGU-3'[190] is used as a control to optimize the experimental parameters.

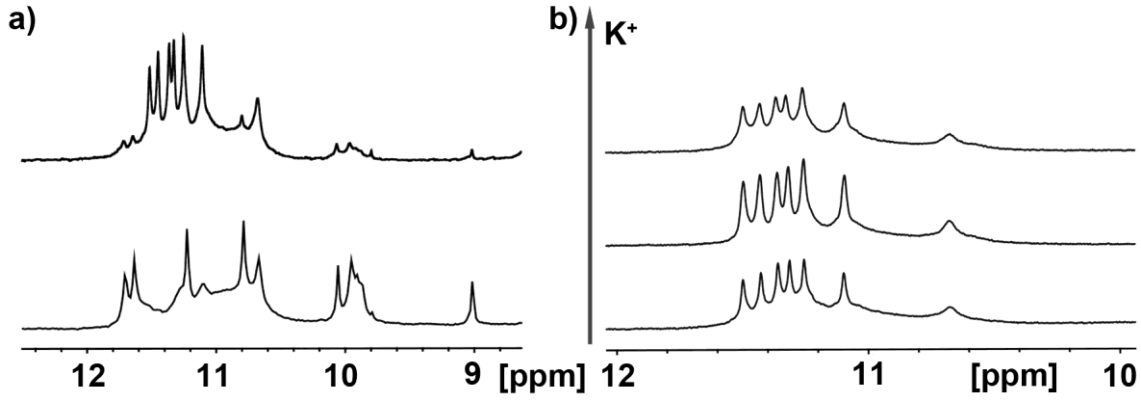


Figure 2.17. ^{1}H NMR of 12-nt TERRA G4, imino proton region. a) 12-nt TERRA 0.14 mM in 90 % H_2O / 10 % D_2O at 25 °C with 1 mM K^+ (top) or no added cations (bottom). Pre-folding or pre-stacking of guanine bases is seen in absence of monovalent cations while addition of K^+ yields six resolved imino peaks; b) 12-nt TERRA 0.5 mM in 1 mM potassium(I) phosphate buffer, pH 7.4, at 25 °C. From bottom to top: 1, 10, 20, and 40 mM K^+ in total.

The spectrum obtained under water-only conditions shows some unresolved peaks in the imino proton region between 10 and 12 ppm (*Fig. 2.17.a*) indicating guanine pre-stacking. Moreover, an extra peak is observed at 9 ppm (*Fig. 2.17.a, bottom*), which might correspond to U:U stacking[275]. Addition of 1 mM K^+ already stabilizes the G4 structure, yielding the same six imino proton peaks reported in the literature, between 10.5 and 11.5 pmm (*Fig. 2.17.a, top*)[185]. Although twelve guanines intervene in G-quartet formation, only half of them are observed because of the two-fold symmetry of the bimolecular G4 structure[195]. An additional peak at around 10.6 ppm is seen herein, which does not correspond to a guanine imino proton.

According to Martadinata *et al.*, the first aromatic proton peak appears at 8.6 ppm (H8 from A8), so it does not seem that the extra peak could be an aromatic peak.

Further titration with K^+ yields a most-resolved spectrum at 20 mM (*Fig. 2.17.b*), while at 40 mM the imino peaks start having an increased baseline signal, which might be due to multimerization effects[195].

Effect of the cation type and concentration

Without addition of any monovalent cations, NRAS and TERRA G4 RNAs in water yield a broad signal around 11 ppm (*Fig. 2.18. and Fig. 2.19.*). The fact that this signal is in the region typical for G4 structures indicates that there is already some pre-stacking of guanine bases (*vide supra*). Subsequently, K^+ cation is added at increasing concentrations. This originates in several peaks being resolved from the initial broad signals (*Fig. 2.18.*). In the case of NRAS (*Fig. 2.18.a*) the spectrum is most resolved in 60 mM K^+, where up to four peaks are detected at 10.3, 10.7, 11, and 11.4 ppm. For TERRA, the optimal concentration is 20 mM K^+, which yields six (albeit not completely resolved) imino peaks at 10.6, 10.65, 10.75, 10.9, 11, and 11.1 ppm (*Fig. 2.18.b*). The fact that the imino peaks are not fully resolved, together with the high background baselines observed (especially in the case of NRAS) give an indication that G4 multimerization might be occurring[117, 190].

NMR spectra are also recorded in presence of Sr^{2+} to follow the effect of divalent cations with strong attraction for RNA. NRAS with 25 equivalents of Sr^{2+} (0.4 mM) yields different spectra in two different measurements (*Fig. 2.19.a and b*). One is well-resolved, with six imino peaks between 11.2 and 11.8 ppm (*Fig. 2.19.a*), while the other shows only a broad peak at around 10.8 ppm (*Fig. 2.19.b*). The sample in *Fig. 2.19.b* corresponds to the same RNA from the sample in *Fig. 2.19.a* after undergoing a desalting treatment by size-exclusion chromatography followed by lyophilizing. Afterwards, the RNA is re-dissolved and Sr^{2+} is added again. One possible explanation for the difference seen between *Fig. 2.19.a* and *Fig. 2.19.b* would be that residual monovalent cations were present before desalting, contributing to higher G4 stabilization. Addition of further excess of Sr^{2+} to the same samples (1000, 2000 or 5000 equivalents, which correspond to 15, 30 or 75 mM Sr^{2+}, not shown here) results in the disappearance of all imino proton signals, presumably because of G4 aggregation (Chapter 3)[274].

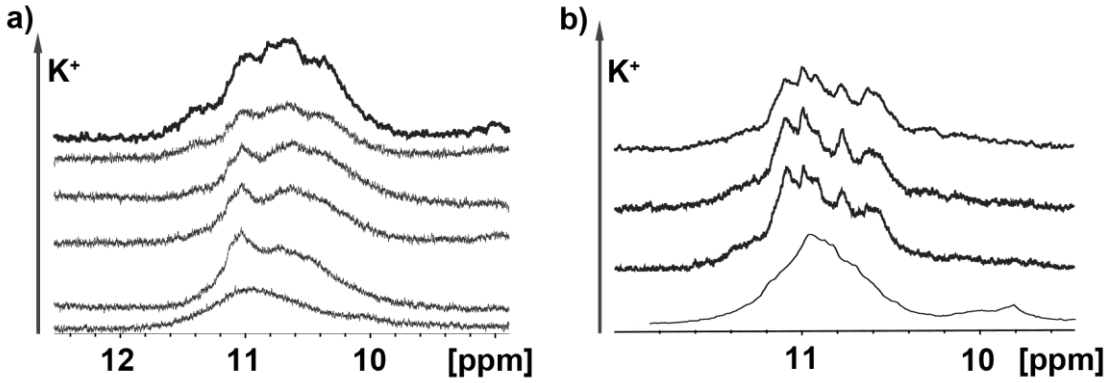


Figure 2.18. ^{1}H NMR of G4 RNAs upon K^+ titration, imino proton region. a) NRAS 0.4 mM in 90 % H_2O / 10 % D_2O at 25 °C. From bottom to top: 0 mM K^+, addition of 1mM potassium phosphate buffer, pH 7.4, and 20, 30, 40, and 60 mM K^+ total. b) TERRA 0.3 mM at 25 °C in 1 mM potassium(I) phosphate buffer, pH 7.4, prepared in 90 % H_2O / 10 % D_2O. From bottom to top: 20, 40, and 60 mM K^+ total.

TERRA with 5 mM Sr^{2+} shows three unresolved proton peaks in the imino region (*Fig. 2.19.c*), although the background baseline is very high, which might be indicative of RNA aggregation. Upon further addition of the divalent cation up to 200 mM Sr^{2+} the imino proton signals disappear almost completely (*Fig. 2.19.d*), which is presumably due to the G4 units being aggregated, as seen for NRAS. The fact that the NRAS spectra appear more unresolved can be attributed to the fact that this sequence contains 13 guanine bases, *i.e.* one more than needed to form a 3-quartet G4, and therefore several G4 conformations might co-exist in solution.

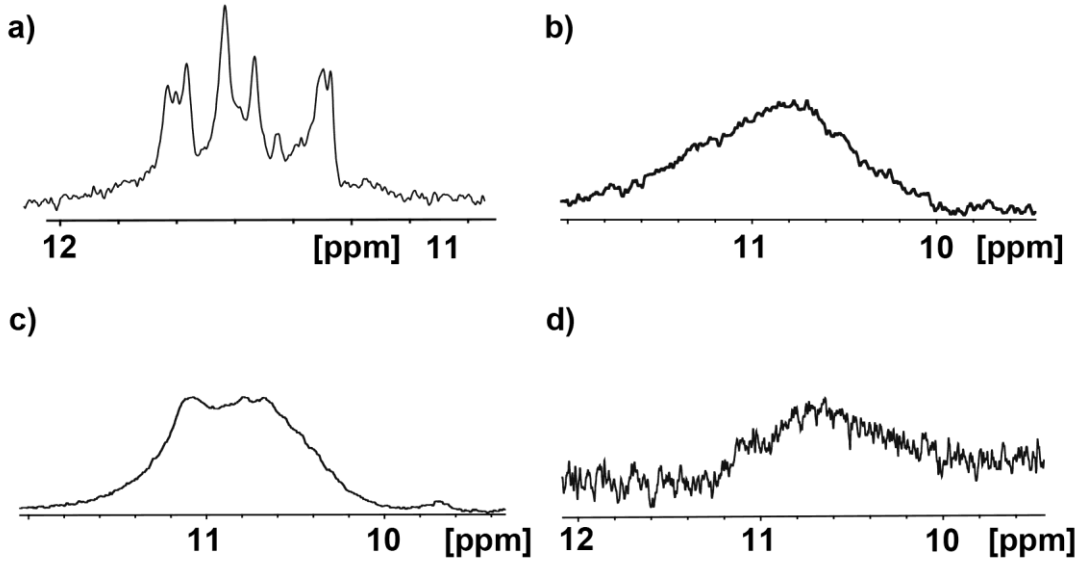


Figure 2.19. ^{1}H NMR of G4 RNAs in Sr^{2+}, imino proton region. a) 15 µM NRAS RNA with 0.4 mM Sr^{2+}, b) the same sample of 15 µM NRAS RNA after desalting and re-addition of 0.4 mM Sr^{2+}, c) 0.2 mM TERRA with 5 mM Sr^{2+}, d) 0.1 mM TERRA RNA with 200 mM Sr^{2+}. All samples are measured in 90 % H_2O / 10 % D_2O at 25 °C.

Effect of the temperature

The effect of temperature on the 1D ^{1}H spectra of the TERRA G4 in K^{+} is examined from 5 to 47 °C. Recording at 5 and 15 °C yields less resolved spectra than the RNA at room temperature (*Fig. 2.20.a*) with less imino proton peaks being observed between 10 and 12 ppm. Better resolution is achieved when the temperature is increased up to 47 °C where five imino proton peaks are present (*Fig. 2.20.b*). The fact that an increase in temperature yields slightly more resolved spectra is in agreement with the hypothesis that the unresolved spectra might be originated by the multimerization of G4 monomers (*Chapter 3*), which seems to be hindered to some extent at higher temperatures.

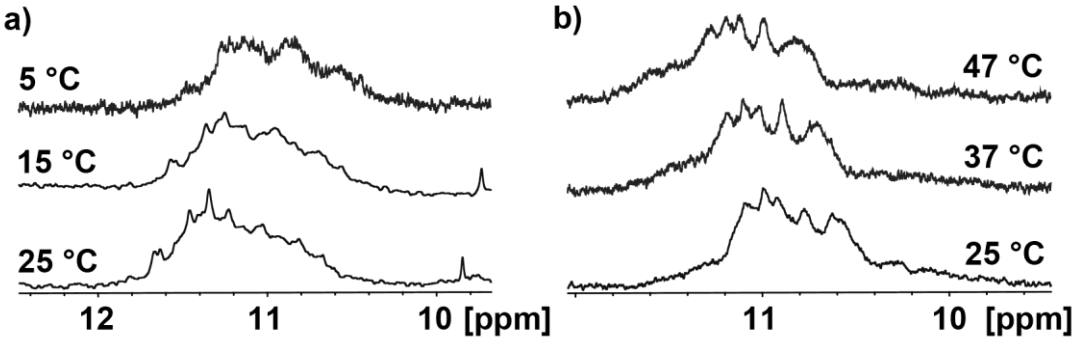


Figure 2.20. Effect of the temperature on TERRA RNA, imino proton region. a) 0.1 mM TERRA RNA in 10 mM phosphate buffer, pH 7.4 and b) 0.3 mM TERRA RNA in 1 mM phosphate buffer, pH 7.4, with 60 mM K^{+}. Both samples measured in 90 % H_2O / 10 % D_2O at 5-47 °C.

Effect of crowding agents

There has been a tendency in recent years to use crowding agents for *in vitro* experiments with nucleic acids in order to better mimic *in cellulo* conditions[276]. These agents, which have a water depletion effect and change the dielectric constant of the solution, include polyethylene glycol (PEG), acetonitrile (ACN), dimethyl sulfoxide (DMSO), and ethanol, among others. When added to htelo G4 DNA in solution, they have been shown to induce conformational transitions and favor the parallel G4 fold to the detriment of the antiparallel structure[215].

Herein, we attempt to use PEG_{200} and deuterated acetonitrile (d_3-ACN) to reduce conformational hetereogeneity of the G4 RNAs. Both NRAS and TERRA yield better resolved spectra upon addition of 40 % (v/v) of either PEG_{200} or ACN (*Fig.2.21.*). In particular, in the presence of PEG_{200} up to seven imino peaks are observed for both sequences indicating that this polymer might stabilize one particular conformation of the G4 RNAs[277,278].

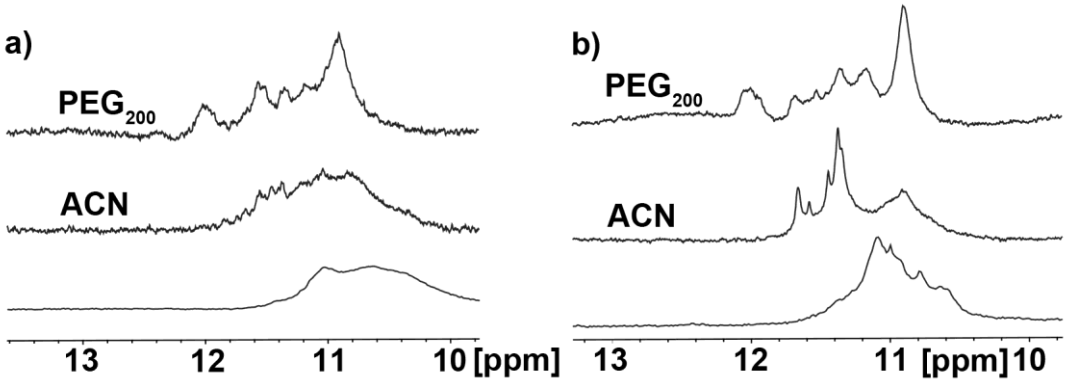


Figure 2.21. ^{1}H NMR of G4 RNAs in crowding agents, imino proton region. a) NRAS 0.1 mM with 20 mM K^+. From bottom to top: no crowding agents, 40 % (v/v) d_3-acetonitrile, and 40 % (v/v) PEG_{200}. b) TERRA 0.1 mM with 20 mM K^+. From bottom to top: no crowding agents, 40 % (v/v) d_3-acetonitrile, and 40 % (v/v) PEG_{200}. Both samples measured in 90 % H_2O / 10 % D_2O at 25 °C.

Effect of a G4-stabilizing ligand

The commercially available 5,10,15,20-tetrakis(1-methyl-4-pyridinio)porphyrin tetra(*p*-toluenesulfonate) compound (TMPyP4) has been reported to have a high affinity for G4 motifs although it lacks specifity[176] (*Chapter 1*). We add 1 equivalent of TMPyP4 to the TERRA G4 to see whether it improves the resolution of the ^{1}H NMR spectrum. In presence of 20 mM K^+ new aromatic peaks between 7 and 9.5 ppm appear (*Fig. 2.24.a vs. b*) with ligand addition. Comparing with the spectrum of the ligand alone (*Fig. 2.24.c*), the aromatic peaks at 7.1 and 7.5 ppm can be assigned to TMPyP4. Additionally, upon ligand addition, the TERRA imino proton peaks between 10 and 12 ppm disappear. These observations indicate that the TMPyP4 compound is indeed interacting with the TERRA G4. However, it seems that the ligand results in the unfolding of the G4, as has already been reported with the MT3-MMP RNA G4[279].

Finally, ^{1}H diffusion NMR (DOSY) was attempted for both RNA G4s in different K^+ concentrations and upon addition of TMPyP4. However, analysis of the data did not yield any consistent diffusion coefficient values (*data not shown*).

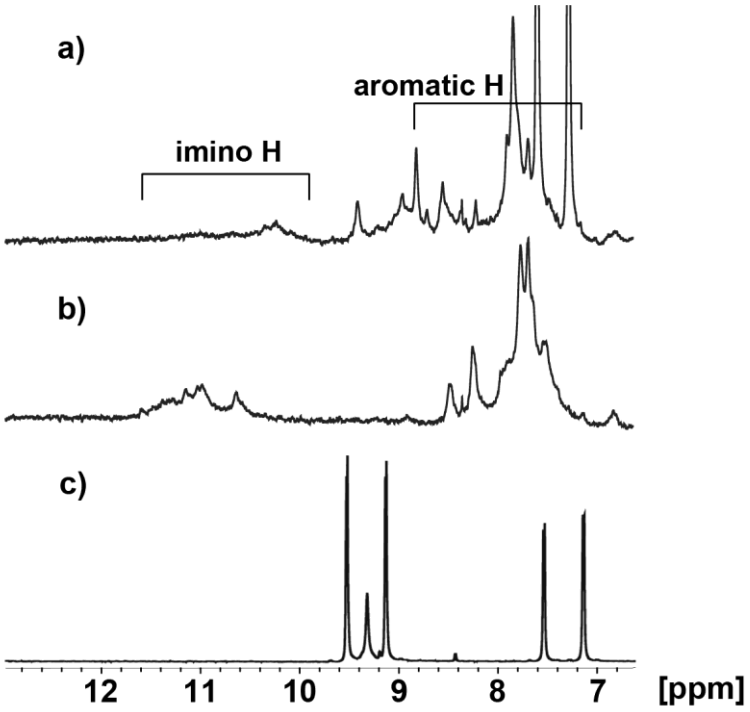


Figure 2.22. ^{1}H NMR of TERRA G4 with addition of the G4 ligand TMPyP4, imino proton region. TERRA 0.075 mM in 90 % H_2O / 10 % D_2O at 25 °C with 20 mM K^+. a) With 1 equivalent of TMPyP4 added and b) without. c) TMPyP4 ligand.

2.2.9 Comparison with the literature on G4 DNA

The folding into parallel G4s, as observed in CD experiments of NRAS and TERRA RNAs (*Figs. 2.3., 2.4. and 2.6.-2.8.*) is consistent with the fact that only the parallel G4 fold is known so far for RNA G-quadruplexes[280]. This is in contrast to G4 DNA, which shows a varied array of topologies[281], exchangeable through changes in temperature or cation conditions[70]. A clear example can be found in the comparison between the telomeric sequences htelo and TERRA. The htelo DNA folds into either parallel, antiparallel or hybrid G4 forms, depending on the identity of the cation present in solution (e.g. Na^+ or K^+) (*Fig. 2.23.b*), the flanking nucleotides, and whether the DNA is in solution or packed in a crystal[121,125,282]. DNA can adopt either the *syn* or the *anti* glycosidic bond angle conformation resulting in an antiparallel and parallel G4, respectively. In contrast, the equivalent TERRA RNA shows only the parallel fold in presence of either K^+ or Na^+ [283] (*Fig. 2.4.*) because RNA prefers the *anti* conformation as a result of the C3-endo sugar pucker, which is favored by the steric constraints imposed by the extra 2'-hydroxyl groups.

In contrast with htelo, dNRAS does not show the same conformational heterogeneity and folds only as a parallel G4 in the presence of either K^+ or Na^+ (*Fig. 2.23.a*), corresponding to the NRAS G4 RNA (*Fig. 2.4.*). Presumably, dNRAS prefers the parallel fold because of its shorter, more stabilizing loops (*Fig. 2.1.*)[120].

In general, considerably higher T_m values are observed for the RNA sequences compared to their DNA counterparts, *i.e.* TERRA vs. its corresponding telomeric DNA sequence (htelo)[227], and NRAS vs. its analogous dNRAS, in all monovalent cation conditions tested herein (*Table 2.3.*). However, ΔT_m's upon monovalent cation addition are comparable, e.g. 12-33 °C upon addition of 20 mM K^+ and 22-26 °C upon addition of 0.1 mM Sr^{2+}. Therefore, the higher RNA T_m values arise from the intrinsic higher stability of G4 RNA in water-only conditions (*Table 2.3.*)[128]. This is rationalized mainly through the presence of an extra OH group in RNA, 2'-OH, which is able to create an extended intramolecular hydrogen bond network, thus further stabilizing the structure[148].

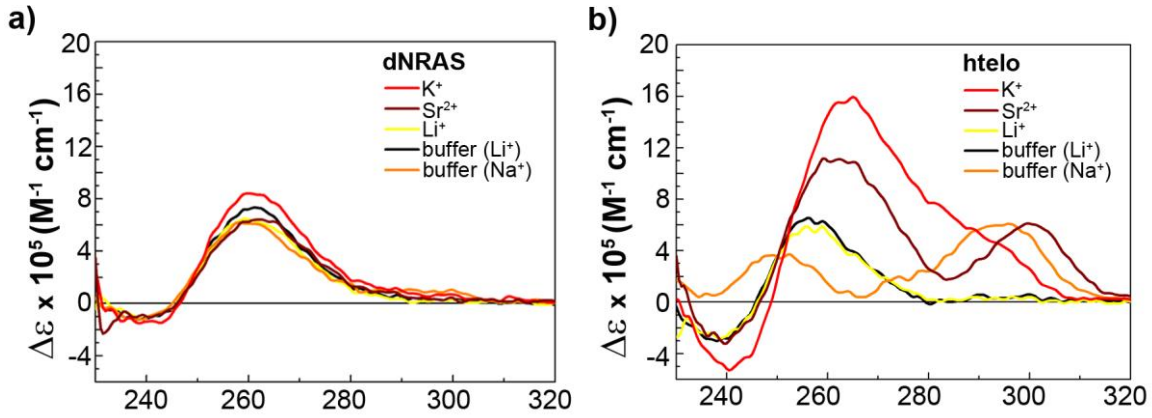


Figure 2.23. CD spectra of G4 DNA in M^{n+}. Effect of K^+, Sr^{2+}, Li^+, and Na^+ on the G4 DNA sequences analogous to the studied RNAs, a) dNRAS and b) htelo, and comparison to the DNAs in both 1 mM sodium(I) MOPS and 10 mM lithium(I) MOPS (both pH 7.4), without further added cations. [DNA] = 10 µM, $[M^+]_{added}$ = 20 mM, $[Sr^{2+}]_{added}$ = 0.1 mM; 10 mM lithium(I) MOPS buffer, pH 7.4, for both M^+ and Sr^{2+} samples.

The general pattern of stability of G4 DNA for monovalent cations has been previously reported to be the following[126]: $K^+ >> Na^+ \geq Rb^+ > Cs^+ \geq Li^+$. Herein, a dramatic stabilization of the G4 RNAs is only observed with K^+, while all other M^+ are undistinguishable by their T_m's. It is significant that no stabilization effect by Na^+ is observed for the two RNA sequences, in contrast to htelo and many other G4 DNAs[227,128]. This high preference for K^+ over Na^+ has been previously shown for similar TERRA sequences as well as other G4 RNAs[128] and might be related to the preference of K^+ for parallel structures. Indeed, the addition of Na^+ shifts the G4 DNAs to an antiparallel conformation, which is not possible in RNA.

The general pattern of stability of G4 DNA in divalent cations has been previously reported[126] as: $Sr^{2+} >> Ba^{2+} > Ca^{2+} > Mg^{2+} > Co^{2+} > Zn^{2+}$. The same pattern of G4 stability is found herein for NRAS; while in the case of TERRA, Ba^{2+}, Ca^{2+} and Mg^{2+} have a similar behavior. The fact that Sr^{2+} is the most stabilizing divalent cation corresponds well to the reported observations for G4 DNA, in which Sr^{2+} gives an even stronger stabilization than K^+ to G4 DNA[231,234].

2.3 Conclusions

In conclusion, both NRAS and TERRA RNAs fold into parallel G4s that are already moderately stable in water (T_m's ~ 50 ºC), albeit with low folded fractions. The choice of buffer and, more specifically, of cationic species contained in the buffer, is shown herein to have a significant impact on G4 RNA folding. Non-stabilizing lithium(I) MOPS is preferred to perform CD, TDS and UV melting experiments.

Some of the tested cations promote G4 folding by unspecific screening of charge repulsions within the phosphate backbone. This is the case of Na^+, Rb^+, Ca^{2+}, Ba^{2+}, Mg^{2+}, and NH_4^+ for NRAS, and of Na^+, Rb^+, Ca^{2+}, and Ba^{2+} for TERRA. Only K^+ and Sr^{2+} significantly increase the stability of both G4 folded structures by binding specifically within the channel cavity[284] (*Fig. 2.27.*), owing to their suitable size and dehydration energies. Moreover, a stronger effect on G4 folding of divalent vs. monovalent cations is observed, which is consistent with higher attraction between a higher positive charge and the RNA polyanion.

Interestingly, no stabilization by Na^+ is observed for the two RNA sequences, in contrast to what has been reported for G4 DNAs. This high predilection for K^+ over Na^+ is presumably related to the preference of parallel-stranded motifs for potassium(I). Moreover, the addition of > 10 mM Sr^{2+} leads to a decrease in G4 CD signals, which is related to suspected G4 multimerization. This is also shown by a higher ΔT_m at high NRAS concentration (> 4 – 10 µM) and will be further explored in *Chapter 3*. Finally, the addition of transition metals (Fe^{2+}, Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, and Mn^{2+}) generally destabilizes the G4 RNAs.

^{1}H NMR spectroscopy yields unresolved spectra in the imino proton region, which further suggests that multimerization is occurring, giving rise to a mixture of species in solution. Indeed, and although RNA concentration independence of T_m confirms that the folding is intramolecular at low G4 concentration, it seems that either upon addition of a higher excess of metal ions or at higher RNA concentration, the G4 monomers are aggregating, especially in the case of NRAS (*Chapter 3*). The fact that this sequence yields less resolved spectra can also partially be attributed to the presence of 13 guanine bases, which might result in several co-existing G4 conformations.

As expected, G4 RNAs are much more stable than the analogous DNAs due to the extra 2'-OH in RNA building further hydrogen bonds within the structure and with the water-based buffers. Remarkably, significant differences exist between the two RNA sequences, e.g. the effect of NH_4^+ and Ba^{2+} on NRAS and the generally higher sensitivity of NRAS vs. TERRA towards cation addition. These differences might potentially be exploited when designing targeting strategies for antitumor-relevant G4 RNAs.

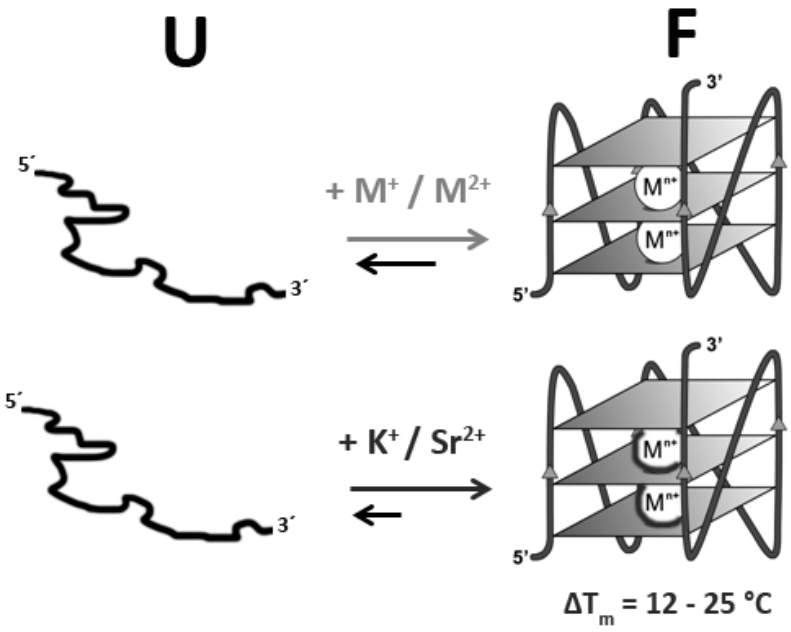


Figure 2.24. Folding of G4 RNAs upon cation addition. Some of the tested cations (Na^+, Rb^+, Ca^{2+}, Ba^{2+}, and Mg^{2+} and NH_4^+ for NRAS) promote G4 RNA folding by unspecific screening of charge repulsions within the phosphate backbone. Only K^+ and Sr^{2+} significantly increase the stability of the G4 folded structures by binding specifically within the channel cavity[279], owing to their suitable size and dehydration energies.

Chapter 3

Multimerization of G4 NRAS RNA in excess of stabilizing cations

3.1 Introduction

RNA G-quadruplexes are often studied *in vitro* as short G-rich sequences (15-25 nt) and have been shown to aggregate by cation-mediated π-stacking of two or more G4 monomers. A previously reported mass spectrometry (MS) study on telomeric sequences showed that G4 RNAs have a higher tendency to multimerize compared to their analogous DNAs[129]. The authors reported that two telomeric RNAs, 12- and 22-nt in length, form dimers by cation-mediated stacking of two parallel G4 units, each consisting of three G-quartets and two ammonium ions[129]. The solution structure, solved by NMR spectroscopy, of a shorter telomeric RNA, the 10-mer 5'-GGGUUAGGGU-3', contains two dimeric three-layer parallel-stranded G4 blocks that stack on each other at their 5' ends, with simultaneous rearrangement of the loops[285].

A few dimeric solution structures have already been solved for G4 DNAs, for example the HIV-1 integrase inhibitor T30177, which forms a stacked parallel-stranded, dimeric G4[286]. The same is true for the minisatellite sequence CEB1, 5'-AGGGGGAGGGAGGGTGG-3', which folds into a parallel-stranded G4 containing three tetrads and shows two subunits stacked at the 5' end[198]. The monomer to dimer transition is observed in PAGE gels between 0.1 and 3 mM K^+ content (*Fig. 1.18.*).

A screening with DNA sequences containing four repeats of three guanines has revealed that sequences with short loops, which favor parallel-stranded G4s, form very stable multimers even at low DNA concentrations as there is no loop impediment to G4 monomer stacking[117]. Moreover, the authors concluded that G4 monomers are favored when flanking sequences are added, whereupon G4 trimers are no longer seen – however, dimers can still be formed.

During the spectroscopic measurements[287] reported in *Chapter 2* some observations suggested that NRAS G4 multimerization might be occurring (e.g. NMR spectra). Therefore, we subsequently perform native polyacrylamide gel electrophoresis (PAGE) together with electrospray ionization mass spectrometry (ESI-MS) measurements to determine the G4 RNA strand stoichiometry of the 18-nt NRAS and 24-nt TERRA RNAs in the presence of an excess of stabilizing cations, namely K^+ and Sr^{2+}.

PAGE gels show possible dimerization of NRAS, which is confirmed by MS analysis. The effect of flanking nucleotides is assessed by elongating the 18-nt NRAS sequence (NRAS18) to 20-nt NRAS (NRAS20), 21-nt NRAS (NRAS21), 22-nt NRAS (NRAS22); and by shortening the 24-nt TERRA (TERRA24) to 21-nt TERRA (TERRA21). The addition of non-guanine bases to the 5' end hinders aggregation of the quadruplexes suggesting that the dimerization proceeds by 5'-5' end stacking. Finally, dynamic light scattering allows the determination of

the hydrodynamic radii and corroborates the tendency towards further multimerization with a higher excess of stabilizing metal ions.

3.2 Results and discussion

3.2.1 Native PAGE experiments show slow-running NRAS G4 multimers

Both 18-nt NRAS (NRAS18) and 24-nt TERRA (TERRA24) contain four repeats of consecutive guanines and are therefore expected to form a G4 by intramolecular folding[120]. Indeed, T_m vs. RNA concentration confirms the intramolecular G4 formation at low RNA and M^{n+} concentrations (*Section 2.2.4 and Fig. 2.2.10.*). However, there are several indications that some further species are present in solution, especially in the case of NRAS (*vide supra, e.g. Fig. 2.2.10.*, NRAS at higher RNA concentration).

Prior to experiments, all DNA and RNA sequences are re-purified by native PAGE electrophoresis followed by two desalting steps (*Methods, Section 5.1.3.*). Denaturing PAGE gels confirm that all DNAs and RNAs are indeed pure with only one band per lane observed in all cases (*Fig. 3.1.*).

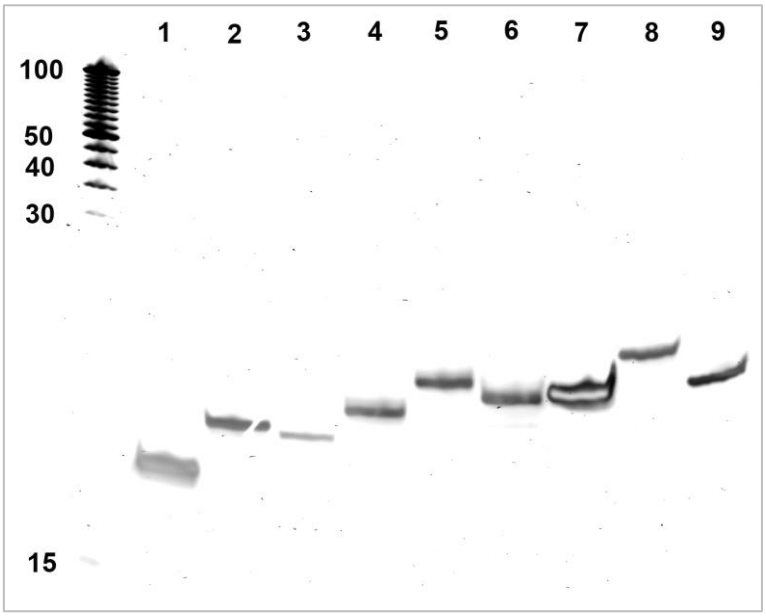


Figure 3.1. Denaturing PAGE showing the purity of the DNA and RNA sequences. 20 % polyacrylamide gel containing urea, run at r.t. in 1 × TBE buffer with 25 pmol RNA/well. Lanes: 1) dNRAS18, 2) NRAS18, 3) tNRAS18 ("t" stands for translated; in contrast to the commercially obtained RNA), 4) NRAS20, 5) NRAS22, 6) NRAS 21, 7) htelo24, 8) TERRA24, 9) TERRA21. A 5 bp dsDNA ladder (commercial) is used as a size marker.

The NRAS and TERRA RNA sequences (NRAS18 and TERRA24) are run in native PAGE gels containing 20 mM K^+ in order to observe the multimerization behavior of their G4 folded states. The corresponding DNA sequences dNRAS and htelo are used for comparison and all bands are subsequently referred to a commercial dsDNA size marker (10-100 bp) loaded in

the first or last lane (*Figs. 3.2. and 3.3.*). Both dNRAS and htelo DNAs appear as single bands with a mobility that corresponds to their size, 18-nt and 24-nt, respectively. They2 can therefore be interpreted as monomeric G4s (*Figs. 3.2. and 3.3.*). The same is true for the TERRA24 RNA (*Fig. 3.2.*). In contrast, NRAS18 shows a mobility twice as slow (*Fig. 3.3.*) than expected for its size (18-nt; compare to denaturing PAGE gel, *Fig. 3.1.*). The NRAS G4 seems to be forming a dimer under these K^+ conditions via stacking of two G4 monomers, although higher aggregates cannot be ruled out (*Fig. 3.6. for an NRAS running without smear*). The faster mobility observed for G4 DNAs compared to the G4 RNAs has already been reported for other G4 sequences[288].

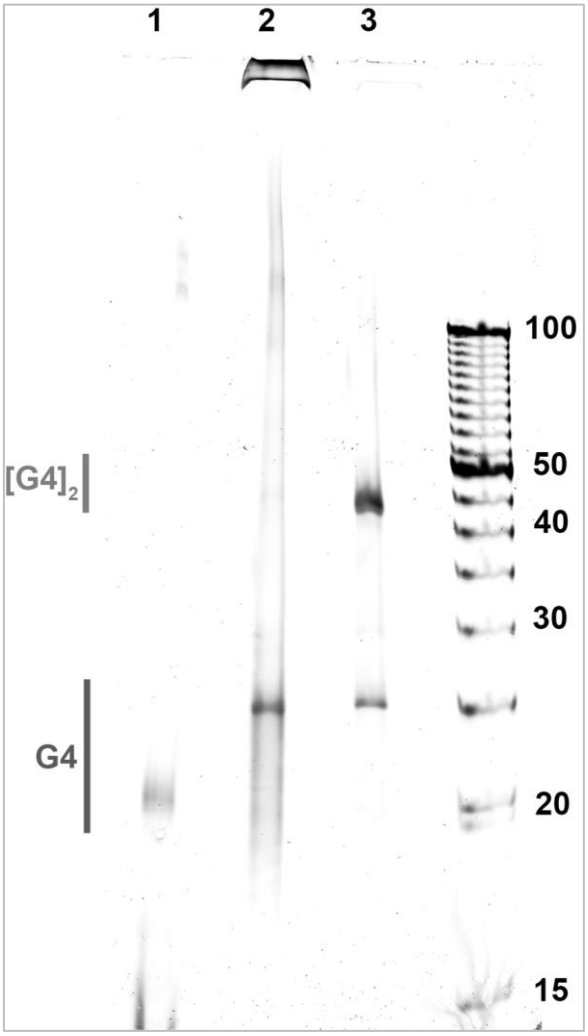


Figure 3.2. Native PAGE of TERRA sequences in 20 mM K^+. 20 % polyacrylamide gel, run at 4 ºC in 1 × TBE buffer with 25 pmol RNA/well. The gel, the running buffer and each of the samples contain 20 mM KCl. Lanes: 1) htelo24, 2) TERRA24, 3) TERRA21. A 5 bp dsDNA ladder is used as a size marker.

It is reported in the literature that formation of G4 multimers is favored by short loops and by the lack of flanking nucleotides[117], which is in agreement with the dimerization being observed for NRAS (5'-GGG, loops 1-2-3 or 2-1-3) and not for TERRA (5'-UUAGGG, loops 3-3-3) in

these conditions. The importance of the 5' flanking nucleotides is further proven by the fact that TERRA24 is monomeric while the related 22nt-TERRA sequence 5'-AGGG(UUAGGG)$_3$-3' has previously been reported to dimerize in 50 mM of NaCl, KCl or NH_4OAc[129].

The fact that only one band is seen for the NRAS RNA (*Fig. 3.6.*) suggests that the dimer species is favored over the monomeric G4 and that there is a preferred G4 end for multimerization as no higher order aggregates are observed. The native PAGE only shows faint bands that might correspond to NRAS trimers or tetramers after storing the K^+ RNA at room temperature for 3 days (*Fig. 3.4.*). It should also be noticed that in all K^+ PAGE experiments a fraction of NRAS18 RNA is found in the gel pockets, which is known to be a sign of larger aggregates[289].

The native PAGE shown in *Fig. 3.3.* is conducted with 20 mM K^+ in the gel matrix, the running buffer, and in each of the samples. Under these conditions, the NRAS monomer is never observed. The faster-running G4 NRAS monomer band is only seen 8when cations are absent both in the gel and in the buffer and K^+ salt is added only to each sample (*Fig. 3.5.*).

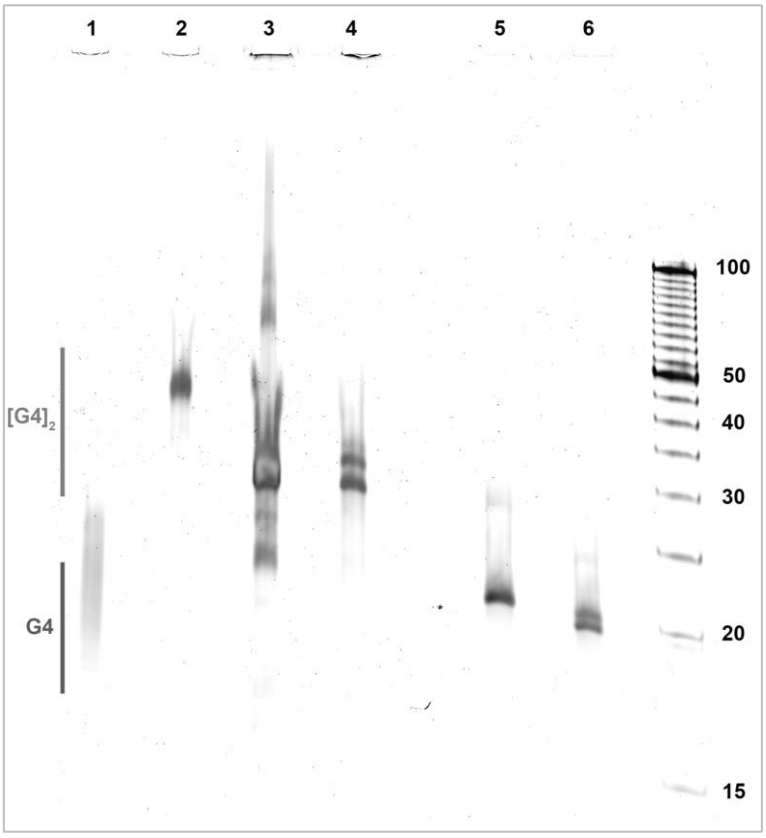


Figure 3.3. Native PAGE of NRAS sequences in 20 mM K^+. 20 % polyacrylamide gel, run at 4 ºC in 1 × TBE buffer with 25 pmol RNA/well. The gel, the running buffer and each of the samples contain 20 mM KCl. Lanes: 1) dNRAS18, 2) NRAS18, 3) tNRAS18, 4) NRAS20, 5) NRAS22, 6) NRAS 21. A 5 bp dsDNA ladder is used as a size marker.

It has been reported that sample preparation can influence the formation of G4 multimers[290]. Therefore, different annealing conditions are tested. After denaturation by heating 5 min at 95 °C, cooling is performed by either a) quick cooling on ice, b) slow cooling to room temperature during 3 h or c) slow cooling to room temperature overnight. No difference is observed on native PAGE gels (*Fig. 3.4.*) and therefore all samples are subsequently treated by quick cooling on ice.

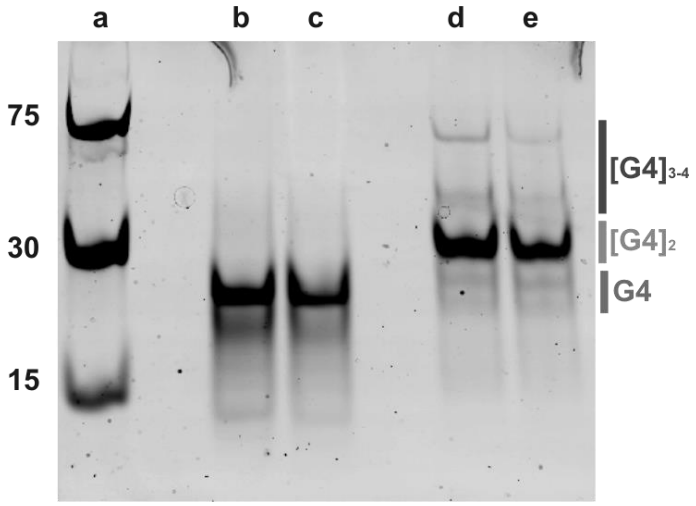


Figure 3.4. Native PAGE in 20 mM K^+ after 3 days. 20 % polyacrylamide gel, run in 1 × TBE buffer with 10 pmol RNA/well. The gel, the running buffer, and each of the samples contain 20 mM K^+. The RNA samples are run after room temperature storage for 3 days. a) Commercial DNA T-ladder, running at 75, 30 and 15 nt; b) TERRA, quick-cooled on ice; c) TERRA, slow-cooled to room temperature; d) NRAS, cooled rapidly on ice; e) NRAS, slow-cooled to room temperature. This gel was run at the Vilar's Lab, Imperial College London.

Gel staining for visualization of the RNA is first tested with GelRed and ethidium bromide (EtBr) (*Appendix, Figs. B.1. and B.2.*). EtBr yields less defined G4 bands, which is probably due to the fact that its interaction with this type of nucleic acid motifs is known to be rather weak[291]. SybrGold is finally chosen due to its higher sensitivity, which allows the RN2A to be observed with less pmol/well, so overloading of the lanes is avoided[292]. Moreover, less background noise is observed upon gel fluorescent scanning with SybrGold than with GelRed.

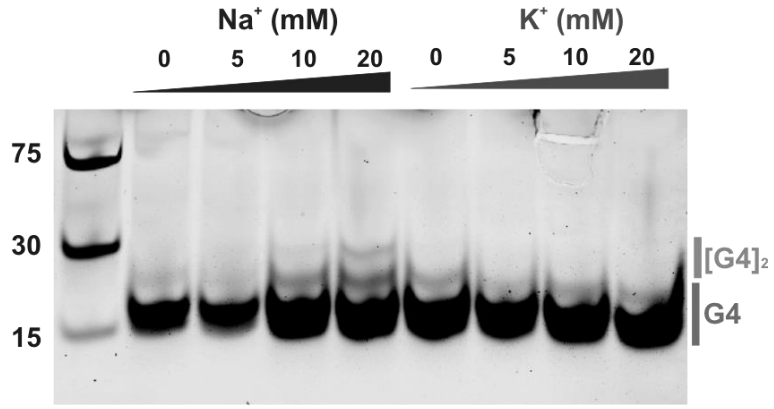


Figure 3.5. Native PAGE without K^+. 20 % polyacrylamide gel of NRAS RNA, run in 1 × TBE buffer with 10 pmol RNA/well. No cations are added to the gel or running buffer, only to each RNA sample. This gel was run at the Vilar's Lab, Imperial College London.

PAGE gels in K^+ show that the commercial NRAS is likely forming a G4 trimer, while the transcribed NRAS18 is prone to dimerization (*Fig. 3.6.*). RNA obtained by *in vitro* transcription with T7 polymerase is isolated with a triphosphate at the 5' end[293], while chemically synthesized sequences of commercial origin have both the 5' and 3' ends as OH[294]. We postulate that the difference in mobility might be due to this difference in phosphorylation at the 5' end, with the triphosphate group yielding steric hindrance to further stacking.

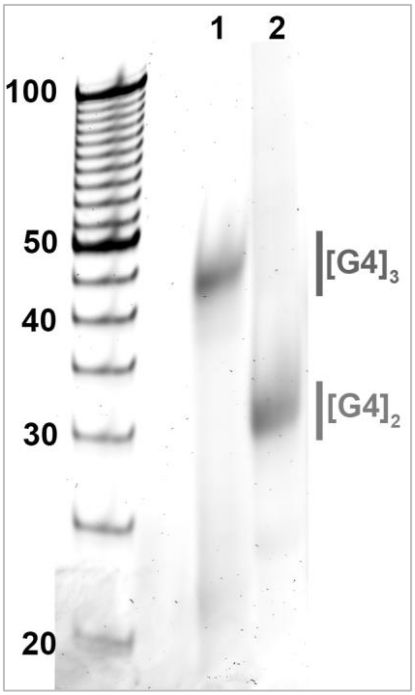


Figure 3.6. Native PAGE of NRAS18 in 20 mM K^+. 20 % polyacrylamid2e gel, run at 4 ºC in 1 × TBE buffer with 25 pmol RNA/well. The gel, the running buffer and each of the samples contain 20 mM KCl. Lanes: 1) NRAS18, commercially obtained and chemically synthesized, 2) tNRAS18, synthesized in house by *in vitro* transcription.

The same NRAS and TERRA RNA sequences (NRAS18 and TERRA24), together with the equivalent DNAs dNRAS and htelo, are run in a native PAGE gel containing 20 mM Sr^{2+}, because of the high affinity of this cation for G4 structures (*Chapter 2*). On2ce more, dNRAS, htelo and TERRA24 all appear as monomers, while both transcribed and chemically synthesized NRAS18 are clearly multimerizing (*Fig. 3.7.*).

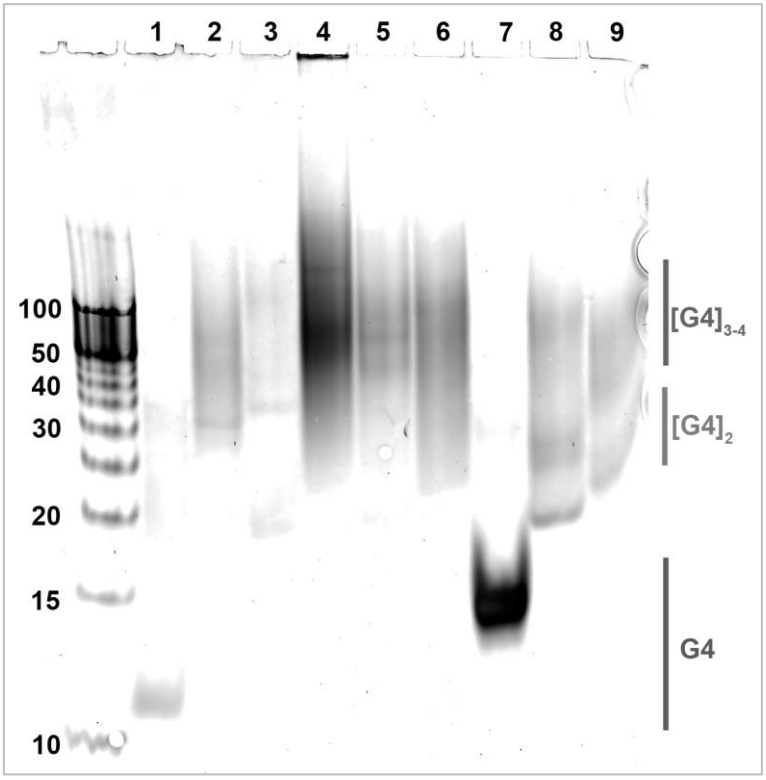


Figure 3.7. Native PAGE in 20 mM Sr^{2+}. 20 % polyacrylamide gel, run at 4 ºC in 1 × TBE buffer with 25 pmol RNA/well. The gel, the running buffer and each of the samples contain 20 mM $SrCl_2$. Lanes: 1) dNRAS18, 2) NRAS18, 3) tNRAS18, 4) NRAS20, 5) NRAS22, 6) NRAS 21, 7) htelo24, 8) TERRA24, 9) TERRA21. A 5 bp dsDNA ladder is used as a size marker.

3.2.2 Extended NRAS sequences

Our working sequence for G4 NRAS, NRAS18, contains only the G4-forming part of the mRNA and presents G bases involved in G-quartets at both the 5' and 3' ends. This RNA folds into a parallel G4 with external loops (*Fig. 2.1.*) and has therefore no loop hindrance or hindrance from flanking nucleotides to the π-stacking of G4 monomers. Multimerization is likely favored in solution as seen by the dimerization observed in native PAGE gels in presence of K^+ or Sr^{2+} (*Fig. 3.3.*). Knowing that the flanking nucleotides have a great impact on the aggregation behavior, the NRAS18 is extended following the natural mRNA sequence and obtaining two new G4 RNA sequences:

- NRAS20, 5'-**U**GGGAGGGGCGGGUCUGGG**U**-3', with one added base at each end

- NRAS22, 5'-**GU**GGGAGGGGCGGGUCUGGG**UG**-3', with two added bases at each end

Circular dichroism (CD) and thermal difference spectra (TDS) confirm the G4 folding of these extended sequences (*Appendix, Fi2gs. B.3.a and B.5.b*). Their melting temperatures are measured in 20 mM K^+ and compared to the T_m of the original NRAS RNA, NRAS18 (*Table 3.1.*). NRAS22 is slightly more stable than NRAS18, with $\Delta T_m = 6\pm1$, while NRAS20 is destabilized, with $\Delta T_m = -10\pm5$ (*Appendix, Fig. B.6.a&b*). However, these changes are presumably not substantial enough to radically affect the G4 behavior.

Native PAGE gels in 20 mM K^+ and 20 mM Sr^{2+} both show multimerization of NRAS20 (*Figs. 3.3. and 3.7.*), which runs as two dimeric conformations in K^+. Therefore, adding one nucleotide at each end of the NRAS G4 does not seem to be sufficient to hinder aggregation. In contrast, addition of two bases at each end yields NRAS22, which appears as a single monomeric band in K^+ native PAGE (*Fig. 3.3.*). This same sequence seems to multimerize in Sr^{2+} (*Fig. 3.7.*), which can be explained by the higher affinity of a divalent vs. a monovalent cation for the same polyanion.

3.2.3 Effect of UUA at the 5' end

In order to explore the effect of the UUA flanking bases at the 5' end, two new sequences are designed. One corresponds to a shortened TERRA, TERRA21, with three consecutive guanines at the 5' end. In parallel, an extended NRAS, NRAS21, has the UUA flanking bases added to the 5' end:

- TERRA21, 5'-**GGG**(UUAGGG)$_3$, without 5' flanking nucleotides

- NRAS21, 5'-**UUA**GGGAGGGGCGGGUCUGGG-3', with three flanking bases at the 5' end

Table 3.1. Melting temperatures of the modified sequences compared to the native TERRA and NRAS G4s. Measured in 20 mM K^+, 10 mM lithium(I) MOPS pH 7.4.

G4 RNA	T_m (°C)
NRAS sequences	
NRAS18	75±1
NRAS20	65±5
NRAS21	79±1
NRAS22	81±1
TERRA sequences	
TERRA24	64±1
TERRA21	68±1

CD and TDS confirm the G4 folding of these modified sequences (*Appendix, Figs. B.3.b, B.4. and B.5.b*). Their melting temperatures are measured in 20 mM K^+ and compared to the T_m of the original NRAS and TERRA RNAs, NRAS18 and TERRA24 (*Table 3.1. and Appendix Fig. B.6.c&d*). Both NRAS21 and TERRA21 result in slight stabilizations, with $\Delta T_m = 4\pm1$.

These changes are presumably not substantial enough to radically change the G4 behavior in solution. Native PAGE gels in 20 K^+ or 20 mM Sr^{2+} are undertaken to observe the effect of these 5' end changes (*Fig. 3.2., 3.3., 3.7.*). In K^+, NRAS21 with UUA at the 5' end does not dimerize while TERRA without these three bases at the 5' end forms a G4 dimer.

This corroborates the strong effect of flanking nucleotides, especially at the 5' end of G4-forming RNAs. The stacking of monomers into dimers is clearly favored at the 5' end and occurs by 5' to 5' stacking as already observed previously in reported G4 NMR structures[198,285,286,295]. The presence of non-guanine nucleotides flanking the G4 at this end acts as a steric hindrance to multimerization as seen in TERRA24, which has three non-guanine bases at the 5' end (*Fig. 3.8.a*). Meanwhile, NRAS18 contains three consecutive G-quartet guanines at the 5' end (*Fig. 3.8.b and c*), resulting in favored aggregation.

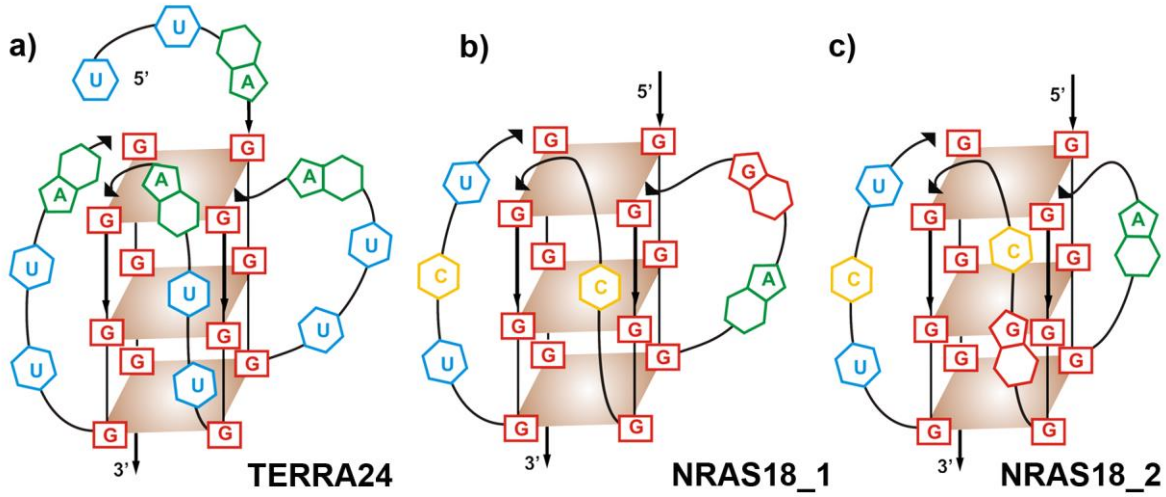


Figure 3.8. Hypothetical structures for TERRA and NRAS G4s. a) TERRA24, 5'-(UUAGGG)$_4$-3'; b) NRAS18, 5'-GGGAGGGGCGGGUCUGGG-3'; c) NRAS18 has two possible conformations, depending on which guanine from the four-guanine run is excluded from the G-quartets.

3.2.4 ESI-MS in NH_4^+ confirms G4 NRAS dimerization

Using a dsDNA size marker as reference for G4 RNA PAGE mobility might lead to some misinterpretations as DNA is being compared with RNA, and double-stranded with G4 structures[296]. Therefore, a complementary technique is required to confirm the multimerization tendency of the G4 DNAs and RNAs. Electrospray ionization mass spectrometry (ESI-MS) is thus performed. The soft ionization conditions allow to preserve non-covalent aggregates in the gas phase[297]. MS in high K^+ concentrations would yield non-resolvable aggregates[298], and therefore the measurements are undertaken in 100 mM NH_4^+ instead.

The MS spectra show that NRAS221, dNRAS18 and TERRA24 are monomeric species (*Fig. 3.9.*). NRAS21 contains little dimer traces, as well as dNRAS18. For TERRA24 the

monomer:dimer ratio is ca. 1:1. On the other hand, tNRAS18, NRAS20 and TERRA21 are mainly dimers, with NRAS20 and TERRA21 showing slight monomeric traces (*Fig. 3.9.*). Native PAGE gels in 100 mM NH_4OAc are performed with all G4 sequences to make sure that the substitution of 20 mM K^+ for 100 mM NH_4^+ does not affect the multimerization tendency (*Fig. 3.10.*). The results are the same as for the K^+ gel and in agreement with the MS data. The only exception is NRAS21, which seems to dimerize in 100 mM NH_4^+, probably due to the higher cation concentration vs. 20 mM K^+. All in all, we can confidently say that NRAS18 forms a G4 dimer in K^+ or NH_4^+ solutions while TERRA24 does not.

MS observations deviate slightly from what is seen in PAGE gels with TERRA24 RNA. In the gels, a single band is positioned as a G4 monomer, while the MS suggests a 1:1 monomer:dimer mixture. This could be explained by the higher RNA concentration used in MS experiments (5 µM vs 2.5 µM in native PAGE), as a higher RNA concentration will likely lead to a larger extent of multimerization[117]. Finally, the commercially obtained NRAS18 shown in the second lane of all gels (*Figs. 3.3., 3.4., 3.7. and 3.10.*) appears as a dimer in MS while its PAGE bands are slower than expected, suggesting trimeric or higher aggregates. This difference might be due to the higher aggregates not being stable enough to be maintained under the ionization conditions and thus separating and yielding dimers[129].

From MS data it is not possible to determine whether the NRAS18 dimer is formed by a cation-mediated stack of two monomers[198,285,286,295] or if it is rather an interlocked fold, in which the strands from each monomer are physically linked[295,299]. However, we postulate a dimerization of intramolecular G4 monomers, and not a dimeric G4 formed directly by two NRAS strands, based on the observed melting temperatures, which are independent from RNA concentration at low RNA concentrations[287] (*Chapter 2*) suggesting intramolecular G4 folding.

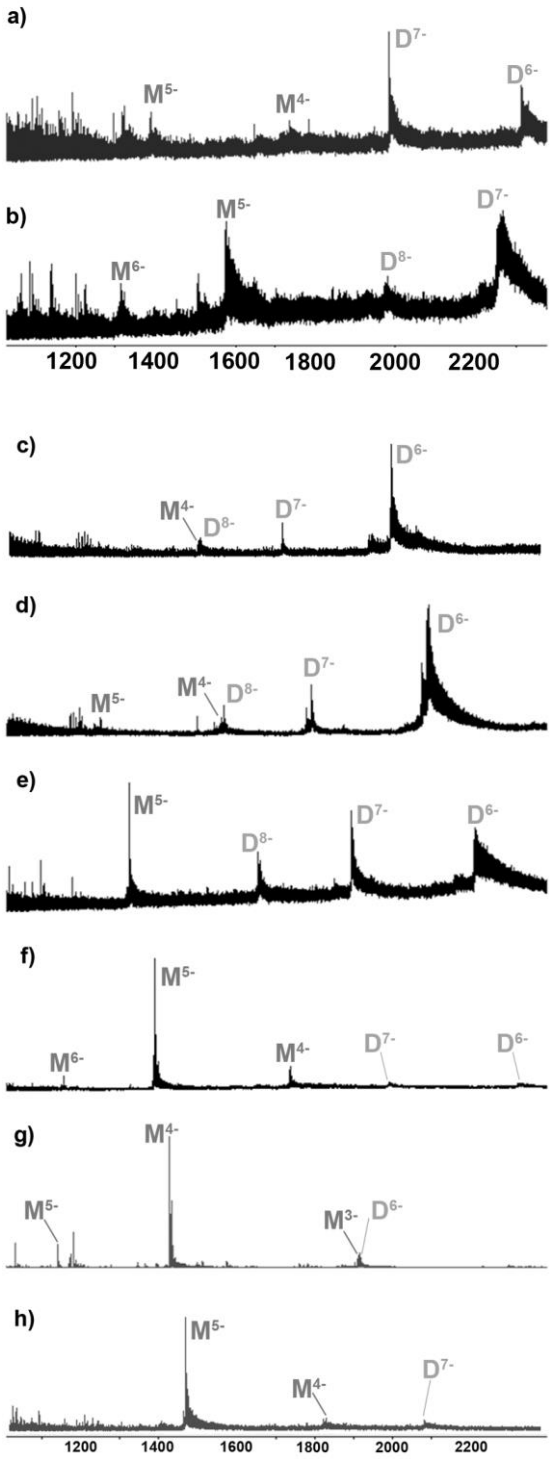


Figure 3.9. ESI-MS of G4 RNAs in 100 2mM NH_4OAc. ESI-MS of annealed G4 DNAs or RNAs at 5 µM in 100 mM NH_4OAc, recorded at the University of Bordeaux by A. Marchand. The peak annotations $[M/D]^{z-}$ indicate the strand stoichiometry, monomer (M) or dimer (D), and the total charge (z). a) TERRA21, MW calcd. 6885.01 g/mol; b) TERRA24, MW calcd. 7823.53 g/mol; c) NRAS18, MW calcd. 5960.52 g/mol; d) tNRAS18, MW calcd. 6196.42 g/mol; e) NRAS20, MW calcd. 6570.84 g/mol; f) NRAS21, MW calcd. 6899.04 g/mol; g) dNRAS, MW calcd. 5700.59 g/mol; h) NRAS22, MW calcd. 7259.24 g/mol.

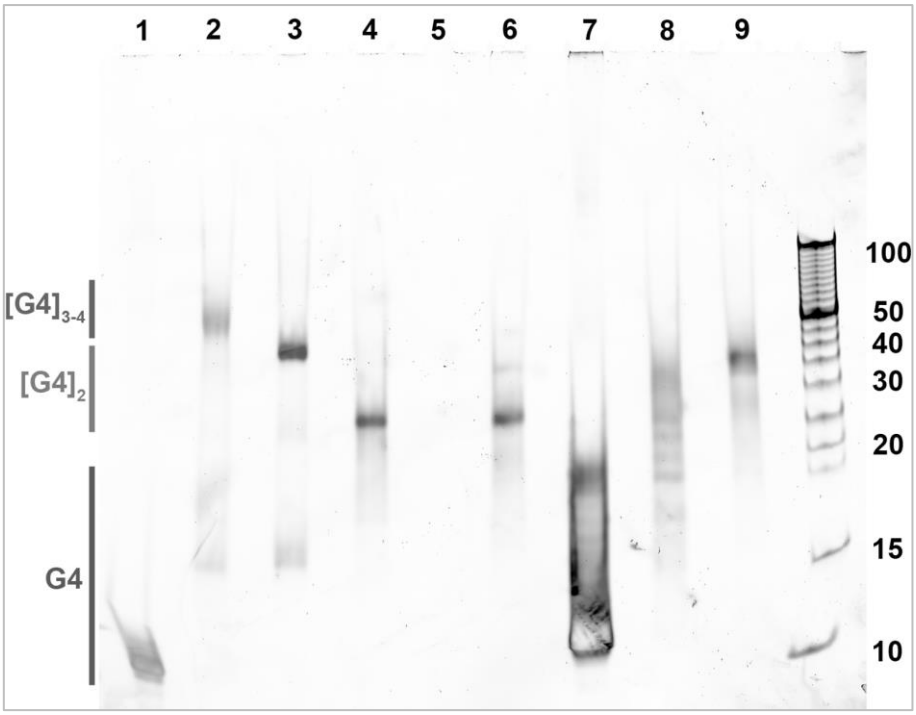


Figure 3.10. Native PAGE in 100 mM2 NH_4^+. 20 % polyacrylamide gel, run at 4 ºC in 1 × TBE buffer with 25 pmol RNA/well. The gel, the running buffer and each of the samples contain 20 mM $SrCl_2$. Lanes: 1) dRNAS18, 2) NRAS18, 3) tNRAS18, 4) NRAS20, 5) NRAS22, 6) NRAS 21, 7) htelo24, 8) TERRA24, 9) TERRA21. A 5 bp dsDNA ladder is used as a size marker.

3.2.5 Determination of the hydrodynamic radii

We determine the size of the NRAS18 and TERRA24 RNA molecules by dynamic light scattering (DLS), which provides a measure of the hydrodynamic radius (r_H) in solution. The aim is to further investigate the cation-mediated multimerization for TERRA and NRAS RNAs in excess of K^+ and Sr^{2+}. The theoretical hydrodynamic radii calculated for NRAS18 and TERRA24 according to *Eq. 1.1. and 1.2.* are listed in *Table 3.2*.

Table 3.2. Theoretical hydrodynamic radii ($r_{H,\ theo}$). Calculated for the NRAS and TERRA G4s by either the spherical (monomer) or symmetric cylinder model (dimer, tetramer).

Spherical model			Symmetric cylinder model					
Monomer			Dimer			Tetramer		
L (nm)	q	r_H (nm)	*L* (nm)	q	r_H (nm)	*L* (nm)	q	r_H (nm)
2.4	1.04	1.20	4.8	2.09	1.85	9.6	4.17	2.57

The NRAS RNA dissolved in water clearly shows that the more K^+ is added the bigger the determined G4 radius becomes (*see Table 3.3.*). Therefore, K^+ addition from 25 to 1000 equivalents (corresponding to 6-240 mM) favors an increasing multimerization of the G4 RNA. After storing the NRAS samples at 4 ºC for 24 h, the r_H values do not change significantly: 2.28±0.10 nm for 25 equiv. K^+, 2.28±0.10 nm for 250 equiv. K^+, and 2.62±0.47 nm for 1000 equiv. K^+ (*compare to Table 3.3.*).

Addition of Sr^{2+} to the NRAS G4 yields higher aggregation than K^+ as seen by bigger determined radii (*Table 3.3.*) and exemplified by the behavior in 250 equivalents (60 mM), which results in precipitation of the RNA, with the white precipitate containing 96 % of the initial NRAS RNA.

Table 3.3. Experimental hydrodynamic radii ($r_{H, exp}$). Measured for 0.24 mM NRAS RNA in water upon addition of either K^+ or Sr^{2+}.

	equivalents K^+			equivalents Sr^{2+}		
	25	250	1000	4	25	250
$r_{H, exp}$	2.16±0.08	2.48±0.06	2.89±0.28	1.31±0.41[a] 3.39±1.26	3.51±0.33[a] 15.51±1.10	–
observation	dimer	multimer	multimer	monomer[a] + multimer	multimers	precipitate aggregates

[a] main population.

TERRA shows less effect of K^+ addition (*Table 3.4.*) as expected from its lower tendency towards multimerization observed in the PAGE and MS experiments (*vide supra*). However, addition of the divalent cation Sr^{2+} has a stronger effect than K^+ as seen by an increase in r_H at 25 equivalents (5 mM) and RNA precipitation appearing at 250 equivalents (50 mM) (*Table 3.4.*).

Table 3.4. Experimental hydrodynamic radii ($r_{H, exp}$). Measured for 0.2 mM TERRA RNA in water upon addition of either K^+ or Sr^{2+}.

	equivalents K^+			equivalents Sr^{2+}		
	25	250	1000	4	25	250
$r_{H, exp}$	1.22±0.12	1.81±0.13	1.98±0.09	1.42±0.20	2.45±0.18	precipitate
observation	monomer	dimer	dimer	monomer	tetramer	aggregates

Dynamic light scattering yields indicative molecule sizes which, however, would need to be confirmed by a complementary method (e.g. diffusion NMR spectroscopy) due to the inherent experimental errors of the measurements. So far it can be concluded from the results presented in this section that NRAS18 has a higher tendency than TERRA24 towards multimerization of G4 monomers, as already seen in native PAGE gels (*Section 3.2.1.*). Moreover, the higher the cation concentration added to the solution, the more the G4 RNAs result into multimerization. Finally, the divalent Sr^{2+} has a stronger effect than the monovalent K^+ towards the favoring of the multimerization, which is again in agreement with Section 3.2.1. Therefore, with this method the different multimerization tendencies of NRAS18 vs. TERRA24 and with K^+ vs. Sr^{2+} were confirmed.

3.2.6 Consequences of G4 NRAS dimerization for *in vitro* experiments

We show that the flanking nucleotides to the G-rich RNA sequence have a great effect on the multimerization tendency. The choice of the exact G4-forming sequence is therefore crucial

and will likely influence *in vitro* observations. During spectroscopic experiments concrete examples of the consequences of G4 NRAS dimerization have been observed (*Chapter 2 and vide infra*), illustrating how important it is to keep this multimerization tendency of short G4 RNAs in mind when designing a new G4 sequence to study *in vitro*.

Isodichroic points

Upon measurement of CD at different temperatures the appearance of an isodichroic point, *i.e.*, a CD value that is the same for all curves, is expected if no RNA folding intermediates are present[300]. That is, if an isodichroic point is observed, as is the case for TERRA24 (or TERRA) in 20 mM K^+ (*Fig. 3.11.b*), this means that there is a two-state equilibrium and that the unfolded RNA at high temperature folds into an intramolecular G4 when the temperature is lowered. In contrast, NRAS18 (or NRAS) in 20 mM K^+ shows no isodichroic points (*Fig. 3.11.a*) and this is indicative of the presence of more than two inter-converting species in solution.

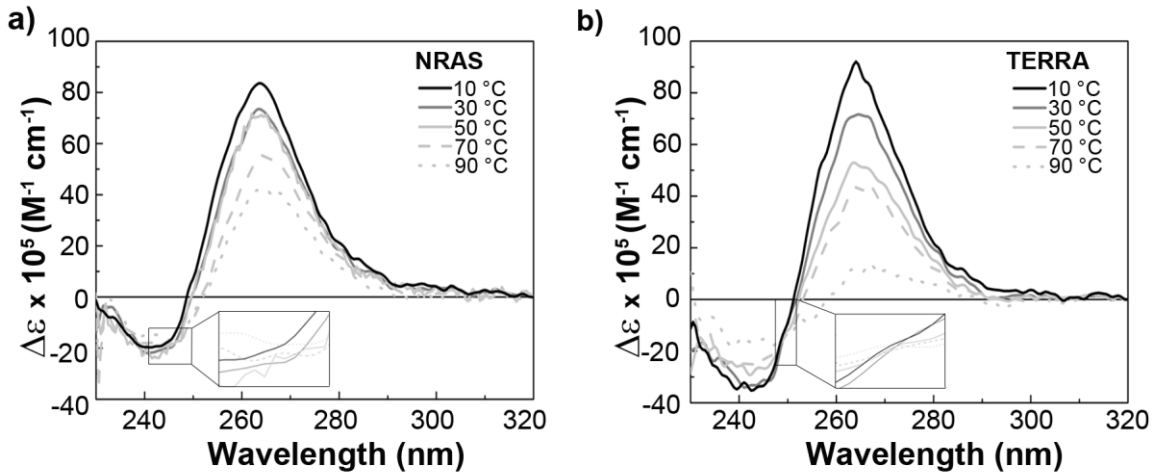


Figure 3.11. Isodichroic points of G4 RNAS. CD spectra of a) NRAS and b) TERRA, measured between 10 and 90 °C in 10 mM TBA MOPS with 20 mM K^+.

Melting curve shape

UV thermal melting experiments in 20 mM of monovalent cations (Li^+, Na^+, K^+, Rb^+, Cs^+, NH_4^+) or 0.1 mM divalent cations (Mg^{2+}, Ca^{2+}, Sr^{2+}, Ba^{2+}) (*Section 2.2.4.*) result in moderate to high hysteresis (*Fig. 3.12. and Appendix, Figs. A.6.-A.9.*) ranging from 2-3 ºC (in Cs^+) to 9-12 ºC (in K^+) for TERRA and from 1-2 ºC (in Rb^+) to 2-15 ºC (in K^+) for NRAS. Hysteresis, defined as the difference between annealing and melting curves in the same sample, indicates a slow denaturation / renaturation process[117]. For both RNA sequences the hysteresis is at its highest in K^+ and two transitions are seen upon heating in these conditions (the melting curve shows two distinct slopes), while only one is observed on cooling (*Fig. 3.12.)*. This behavior has been previously described as indicative of multimer formation[117].2

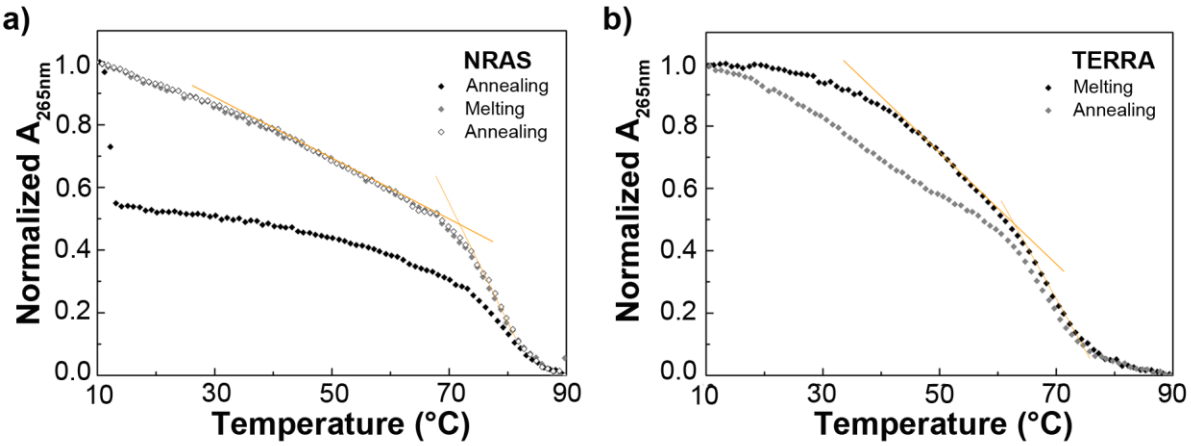


Figure 3.12. Melting curves in 20 mM K^+. Melting and annealing profiles obtained with a) NRAS and b) TERRA, in 10 mM lithium(I) MOPS, pH 7.4, with 20 mM K^+. The two slopes observed for the melting profiles are indicated in orange.

Thermodynamic parameters

Analysis of the melting curves obtained by UV (*Chapter 2*) with the van't Hoff representation (ln*K* vs. 1/*T*) allows to determine the thermodynamic parameters of the G4s for all tested conditions[192]. However, the obtained ΔH_{VH}, ΔS_{VH} and ΔG_{VH} only correctly apply to the system at hand if two states are present in solution (unfolded and folded RNA) and if the ΔH^0 is temperature-independent (*i.e.*, no change in heat capacity occurs). If one of these two conditions is not fulfilled, the thermodynamic parameters will not be reliable.

Table 3.5. Thermodynamic parameters calculated for the NRAS and TERRA sequences in monovalent cations. The measurement conditions were 10 mM lithium(I) MOPS buffer, pH 7.4, and 20 mM of the corresponding M^+.

RNA	T_m (ºC)	ΔH_{VH} (kJ/mol)	ΔS_{VH} (kJ/mol·K)	ΔG_{VH} (kJ/mol) at 37 ºC
NRAS	50±1	− 127±25	− 0.394±0.079	− 4.9±0.6
NRAS-Li	49±1	− 129±20	− 0.400±0.063	− 4.7±0.7
NRAS-Na	50±1	− 165±24	− 0.510±0.074	− 6.6±1.4
NRAS-K	75±1	− 81±28	− 0.228±0.078	− 9.7±3.5
NRAS-Rb	53±1	− 207±38	− 0.636±0.155	− 9.8±2.1
NRAS-Cs	47±1	− 173±67	− 0.536±0.207	− 6.2±3.3
NRAS-NH4	47±1	− 126±15	− 0.392±0.048	− 4.0±1.3
TERRA	52±1	− 166±13	− 0.514±0.042	− 7.4±0.4
TERRA-Li	53±1	− 168±36	− 0.515±0.110	− 8.1±1.8
TERRA-Na	53±1	− 179±24	− 0.550±0.075	− 8.5±1.1
TERRA-K	64±1	− 77±36	− 0.229±0.109	− 6.4±2.3
TERRA-Rb	52±1	− 226±38	− 0.698±0.117	− 10.0±2.3
TERRA-Cs	51±1	− 163±12	− 0.502±0.037	− 6.9±0.9
TERRA-NH4	51±1	− 153±21	− 0.474±0.065	− 6.2±1.0

The thermodynamic parameters obtained by van't Hoff analysis of the NRAS and TERRA in mono- and divalent cations (20 mM and 0.1 mM, respectively) are shown in *Table 3.5.* and *Table 3.6*. Comparison with published values of both G4 RNAs in 5 mM K^+ (*Table 3.7.*) indicates that ΔH_{VH} is possibly overestimated in this work. Indeed, the reported ΔH_{VH} are lower than the ones found herein. As a direct consequence, ΔG_{VH} calculated at 37 ºC (the physiological temperature) is also overestimated and might not represent the real free energy value for the measured cationic conditions. These are all indications that more than two species might be present in solution, which is consistent with the multimerization behavior reported above.

Table 3.6. Thermodynamic parameters calculated for the NRAS and TERRA sequences in divalent cations. The measurement conditions were 10 mM lithium(I) MOPS buffer, pH 7.4, and 0.1 mM of the corresponding M^{2+}.

RNA	T_m (ºC)	ΔH_{VH} (kJ/mol)	ΔS_{VH} (kJ/mol·K)	ΔG_{VH} (kJ/mol) at 37 ºC
NRAS-Mg	48±1	− 126±19	− 0.393±0.060	− 4.3±0.4
NRAS-Ca	54±1	− 147±6	− 0.450±0.019	− 7.7±0.4
NRAS-Sr	75±1	–	–	–
NRAS-Ba	58±4	− 183±10	− 0.551±0.026	− 11.6±2.4
TERRA-Mg	53±1	− 168±36	− 0.515±0.110	− 8.1±1.8
TERRA-Ca	55±2	− 133±17	− 0.406±0.054	− 7.1±0.8
TERRA-Sr	74±2	–	–	–
TERRA-Ba	54±1	− 151±11	− 0.462±0.033	− 7.8±1.2

Table 3.7. Literature values for G4 RNA thermodynamic parameters. Calculated by the authors from melting curves measured in 10 mM lithium(I) cacodylate, pH 7.0, with 5 mM K^+. The same 18-nt NRAS sequence as in this work was used, while the TERRA sequence was 5'-GGG(UUAGGG)$_3$-3' [120].

	$T_{m(VH)}/T_m$ (ºC)±1	ΔH_{VH} (kJ/mol)	ΔS_{VH} (kJ/mol·K)	ΔG_{VH} (kJ/mol) at 37°C
NRAS	69 / 68	− 244±4	− 0.713±0.012	− 22.7±0.5
TERRA	57 / 58	− 265±5	− 0.799±0.016	− 16.3±0.3

CD and melting in excess of M^{2+}

In CD measurements, addition of Sr^{2+} during titrations with this divalent cation results in an unexpected signal decrease at 265 nm (*Section 2.2.6.*). In parallel, additions of 20 mM of either Mg^{2+}, Ca^{2+}, Sr^{2+} or Ba^{2+} to either TERRA or NRAS RNAs yield a distorted CD signature, not typical of a parallel G-quadruplex[287] (*Section 2.2.3.).*

Imino proton region in 1D ^{1}H NMR

^{1}H NMR spectra of both RNA sequences, in both K^+ and Sr^{2+}, show unresolved multimeric peaks in the imino proton region[190] (*Figs. 2.18.-2.25.*), which are also a sign of G4 multimerization, and hinder further structural investigations by NMR spectroscopy.

3.3 Conclusions

Herein we confirm that G4 RNA multimerization is highly sequence-dependent and that the 5' flanking bases are crucial for the aggregation tendency. A 24nt-TERRA RNA containing UUA bases at its 5' end is monomeric even in excess of K^+ while a 18nt-NRAS RNA without any flanking nucleotides at the 5' end dimerizes in excess of K^+, NH_4^+, and Sr^{2+}. When exchanging the UUA end from TERRA to NRAS the aggregation behavior is reversed. ESI-MS performed with NH_4^+ under soft ionization conditions proves to be a good method to confirm the strand stoichiometry and all results are in agreement with the bands observed in native PAGE gels. Structural data on the NRAS G4 would be needed to establish the nature of the dimer, *i.e.* whether it is a simple stacking of G4 monomers or whether some strand interlocking occurring. Finally, it is important to keep this multimerization propensity in mind when performing *in vitro* experiments with short (15-25 nt), G-rich RNA sequences especially if lacking in flanking nucleotides at the 5' end. No reports of G4 multimerization *in vivo* are known so far; however, the low RNA concentrations in the cell, together with the presence of only one G4 sequence per mRNA, make aggregation unlikely in that context.

Chapter 4

Single-molecule measurements on the NRAS G-quadruplex

4.1 Introduction

4.1.1 smFRET of DNA G-quadruplexes

Single-molecule FRET has been proven useful to study the conformational heterogeneity and real-time dynamics of G-rich DNA sequences folding into intramolecular quadruplex structures[221,301]. As a general strategy, a 5' acceptor-labeled G4-forming sequence is followed by a double-stranded overhang containing the donor fluorophore as depicted in *Fig. 4.1.* (acceptor: Cy5; donor: TMR or Cy3)[302]. In some cases, one of the two strands is biotinylated at the 3' end for surface immobilization. Several human G4 DNAs have been studied by this technique including the telomeric sequence (htelo) as well as the c-myc and c-kit promoters. However, there are to date no published smFRET studies of G4 RNAs.

The first smFRET experiments of G4 DNA were performed with the telomeric G4-forming sequence (htelo). Under physiological conditions *in vitro* (37 °C, 140 mM K^+), htelo was shown to be folded as a G4 with comparable populations of parallel and antiparallel conformations[137]. The study of the equilibrium between both G4 conformations is important in order to assess the possible interaction with proteins or small molecules related to the role of htelo in regulating telomere length *in vivo* (*Chapter 1*)[144].

It has been shown for several htelo constructs (*Table 4.1.*) that both the G4 DNA folding from ssDNA and the equilibrium between the parallel and antiparallel folds depend on the identity and concentration of monovalent cations present in the experiment, agreeing with the cation effects reported for G4 DNAs (*Chapter 2*). For example, Lee *et al.* showed three FRET states at 2 mM K^+: low (~ 0.3, attributed to ssDNA), intermediate (~ 0.6, attributed to a parallel G4), and high (~ 0.8, attributed to a antiparallel G4)[137]. The low-FRET conformation was the only one appearing without monovalent ions, and its population decreased as K^+ concentration increased; higher concentrations of Na^+ were required for the same effect. All three FRET states were also temperature-dependent in this study.

Monovalent cations not only affect G4 folding but also the unfolding kinetics as seen with an *O. nova* telomeric G4 DNA when comparing K^+, Na^+, and Li^+ conditions[266]. The unfolding rates in the presence of Li^+ were approximately 1000-fold higher than in the presence of K^+, hinting to the lower stabilization capability of lithium(I) vs. potassium(I) (*Chapter 2*).

Table 4.1. Literature FRET values measured for G4 DNAs. Summary of DNA sequences, cationic conditions, and smFRET states reported in the literature for different G4 DNA constructs.

Construct	Sequence (5'-3')	Conditions	FRET states	Attributed
htelo[302]	Cy5-$\underline{G_3}$TTA$\underline{G_3}$TTA$\underline{G_3}$TTA$\underline{G_3}$... AGAGGTA$_4$GGATAATGGC... CACG GTGCGGACGGC	10 mM Na^+/K^+ 100 mM Na^+/K^+ 10/100 mM Na^+/K^+	0.35 0.52 0.85	parallel G4 parallel G4 antiparallel G4
htelo[137]	Cy5-($\underline{G_3}$TTA)$_3\underline{G_3}$AGAGGT... A$_4$GGATAATGGCCACGGT... GCG-B[a]	2 mM K^+ no M^+	~0.3 ~0.6 ~0.8 ~0.3	ssDNA parallel G4 antiparallel G4 ssDNA
O. nova[266]	Cy5-$\underline{G_4}$T$_4\underline{G_4}$T$_4\underline{G_4}$T$_4\underline{G_4}$AGA... GGTA$_4$GGATAATGGCCA-B	10 mM K^+ 100 mM Na^+ 10 mM Li^+ 30 mM Li^+ 300 mM Li^+	0.2 0.5 0.76 0.65 0.2 0.75 0.64	ssDNA parallel G4 antiparallel G4 antiparallel G4 ssDNA antiparallel G4 antiparallel G4
c-myc[303]	TCTGCT$_3$G$_3$AAC$_3$G$_3$AG$_4$CG… CTT*AT$\underline{G_4}$A$\underline{G_3}$T$\underline{G_4}$A$\underline{G_3}$T$\underline{G_4}$… AAGGT$\underline{G_4}$AGGAGACTCAG… CCG$_3$CAGCCGAGCACTCT... AGCTCTAG	100 mM K^+	~ 0.6 ~ 0.9	parallel G4 various parallel G4
kit-1[304,305]	GAGCAG$_4$CCAGACGCCG… CCG$_3$AAGAAGCGAGAC*C… CG$_3$CG$_3$CGCGAG$_3$AG$_4$AG… GCGAGGAG$_4$CGTGGCCG… GCGCGCAGAG$_3$AG$_3$C	100 mM Li^+ 100 mM Li^+ 100 mM Li^+	~ 0.2 ~ 0.5 ~ 0.7	unfolded partially folded G4 parallel G4
kit-2 [304,305]	as kit-1, different complementary strand	100 mM K^+	~ 0.2 ~ 0.4 ~ 0.7 ~ 0.85	unfolded partially folded G4 parallel G4 parallel G4
c-kit[306]	Cy5-A$\underline{G_3}$A$\underline{G_3}$CGCT$\underline{G_3}$AGG… A$\underline{G_3}$AGAGGTA$_4$GGATAAT… GGCCACGGTGCGGACGGC & Cy5-$\underline{G_3}$C$\underline{G_3}$CGCGA$\underline{G_3}$A… $\underline{G_4}$AG-idem	100 mM K^+ 100 mM Na^+	~ 0.9 0.2-0.8	parallel G4 folding intermediates
3-quartet G4 DNAs[307]	TGGCGACGGCAGCGAG… GC$\underline{G_3}$T$\underline{G_4}$A$\underline{G_3}$T$\underline{G_3}$-Cy3 & 22 more	100 mM K^+	~ 0.55 ~ 0.70 ~ 0.20-0.30	parallel G4 antiparallel G4 unfolded DNA
htelo[308]	GCGTGGCACCGGTAATA… GGA$_4$TGGAGA*T*A$\underline{G_3}$TTA$\underline{G_3}$… TTA$\underline{G_3}$TTA$\underline{G_3}$ & Tel23, with *T*	20/100 mM KCl absence of added salts 100 mM KCl, refolding	~ 0.44 0.72 0.23 0.72 0.50	antiparallel G4 parallel G4 unfolded DNA parallel G4 antiparallel G4
htelo + 2 POT1 [309]	[TTA$\underline{GGG}$]$_4$	150 mM Na^+	0.8 0.6 0.45 0.3 0.2	
htelo + RecA[310]	Cy3-TA$\underline{G_3}$-(TTA$\underline{G_3}$)$_3$-TT-Cy5	100 mM K^+	0.32 0.69 0.86	unfolded DNA folded DNA folded DNA

		with RecA	0.29 0.72	RecA@ssDNA folded DNA
G4 DNA + Pif1 helicase[311]	GGTTACAAGGTTACGTGG... TTACACGGTTACAGGTTA... CAGGTTACA$\underline{G_6}$TTACAGG... TTACGGTTACAGGTTACA... $\underline{G_4}$TTACACAGGTTACAGG... TTACGGTTACA$\underline{G_5}$TTACA... $\underline{G_4}$TTAC$\underline{G_3}$TTACAGGTTAC... A$\underline{G_3}$TTCCACGGTTACGGT... TACA$\underline{G_4}$TTACA$\underline{G_3}$TTACAG$_4$... TACGGTTACAGGTTACAG... GTTACAGGTACGGTTACA... GGTTACA$\underline{G_3}$TTACA$\underline{G_3}$TTA... CAGTTACAGGTTACAGGT... TACA$\underline{G_3}$TTACGGTTACA... $\underline{G_3}$TTACGGTTACA$\underline{G_3}$ttacg... GTTACAGG	5 mM Mg^{2+} & 60 mM K^+	0.77 0.64 0.50 0.32	G4 G-triplex G-hairpin unfolded DNA
3-quartet-G4 DNA + BLM helicase[312]	TGGACCAGACCTAGCAG... CTAT$\underline{G_5}$AGCT$\underline{G_4}$AAGGT$\underline{G_3}$... AATGTGA	100 mM K^+ & 5 mM Mg^{2+}	0.32 0.89 0.72	unfolded DNA antiparallel G4 parallel G4
htelo + BLM helicase[313]	Cy3-TT$\underline{G_3}$TTA$\underline{G_3}$TTA$\underline{G_3}$TTA… $\underline{G_3}$(12T)TGGCGACGGCAG…CG AGGC	150 mM K^+	0.46 0.62 0.26	 BLM-bound unfolded G4
htelo + WRN helicase[314]	($\underline{GGG}$TTA)$_3\underline{GGG}$	50 Mm Na^+	0.2 0.70	unfolded G4 folded G4

[a] B = biotin

Moreover, Long and Stone[308] show that the G4 folding protocol also influences the folded htelo conformations. After thermal annealing in 100 mM or 20 mM KCl, the authors observed two spontaneous FRET populations: one broad at ~ 0.44 that possibly included two different peaks, ~ 0.42 and 0.50, and one narrower at 0.72. In absence of added salts the unfolded DNA yielded a FRET signal at 0.23. Subsequently, *in situ* refolding was performed by adding buffer with 100 mM KCl at room temperature and incubating for 10 minutes. After this time, the smFRET histogram showed one peak at 0.72 (attributed to a parallel G4) and another at 0.50 (attributed to an antiparallel G4). After longer incubations (up to 30 h) more stable mid-FRET conformations appeared, corresponding to hybrid G4s, and the histogram resembled again the one obtained after thermal annealing. The authors interpreted the 0.50 FRET state (antiparallel G4) as a kinetic trap, which increases at higher KCl concentrations. Indeed, it seemed that a high temperature or a low ionic strength facilitated a faster reaching of the thermodynamically favored G4 conformations from the unfolded state.

The G4-forming DNA sequences from human gene promoters are also being researched as possible drug targets for future anticancer strategies[147] (*Chapter 1*). A natural c-myc sequence showed a kinetically stable medium FRET subpopulation at ~ 0.6 in 100 mM K^+ attributed to

a parallel G4[303]. This conformation was ca. 100 times slower to unfold upon complementary strand addition compared to the high-FRET population at ~ 0.9, attributed to a different parallel fold. The presence of a G4 form that is slow to unfold might be of biological relevance and have a role in c-myc silencing, as it would allow *in vivo* interaction with targeting ligands[315].

Different sequences from the native human c-kit promoter have also been studied. Two constructs measured by Shirude *et al.* showed the folding of a parallel G4 from the unfolded DNA, going through a partially folded form. For both constructs, less than 1 % of the molecules showed conformational dynamics during an observation time of ca. 30 min[305]. This lack of dynamics contrasts with htelo and suggests that the dynamic behavior of intramolecular G4-folded conformations in DNA is sequence-specific.

In a second study, two different G4-forming sequences derived from c-kit were also shown to form parallel G4s by CD[306] and presented only one high-FRET species (~ 0.9) in 100 mM K^+, which decreased in the absence of K^+. Interestingly, the calculated half-life of both G4s was in the order of 10^4s, which means that the structures might be folded long enough to allow *in vivo* interaction with small molecules that might act as drug-like ligands (*Chapter 1*).

In comparison with the human promoter sequences c-myc and c-kit, the human telomeric sequence htelo is definitely more heterogeneous in conformation and shows a parallel-to-antiparallel equilibrium that depends on cation type and concentration as well as on temperature. In contrast, both c-myc and c-kit yield parallel-only folds with enough kinetic stability to potentially allow *in vivo* targeting with small molecule ligands.

Not only the nucleotide sequence has an influence on conformation distribution and molecular dynamics, but other aspects such as the differences in loop regions also have an impact. Thus, a kinetic analysis of several intramolecular 3-quartet G4 DNAs showed the effect of loop composition on G4 folding dynamics[307]. Longer loop lengths promoted the antiparallel G4 and unfolded conformations while 1-nt loops favored parallel G4s even when other loops were as long as 9 nt. No G4 formation was observed in absence of cations in 10 mM Tris-HCl, pH 7.5. The shorter the loops, the faster the observed folding kinetics in KCl. Interestingly, the wild-type htelo sequence with TTA loops folded faster than a modified sequence with TAA[307], which might indicate that the wild type htelo DNA has been selected for G4 folding *in vivo*.

In 2012, the first smFRET study of G4 DNA in the presence of binding proteins was reported in 150 mM Na^+ solution[309]. A stepwise decrease in FRET states (0.8 → 0.6 → 0.45 → 0.3 → 0.2) indicated a sequential unfolding of the htelo G4 [TTA<u>GGG</u>]$_4$ by the protein known as protection of telomeres 1 (POT1)[309]. Upon addition of a POT1-TPP1N[316] complex (TPP1N stands for Tripeptidyl Peptidase 1), a fluctuation around mid- to low-FRET values implied a continuous unfolding and refolding of the G4 due to a sliding motion of the protein complex.

This motion might be biologically important for telomerase activity and relevant for telomeric G4 targeting with antitumor purposes[142].

During the last three years, several reports of G4 DNA-protein interactions studied by smFRET have been published. The effect of the cation type and concentration, the presence of adenosine tryphosphate (ATP) and the design of the DNA construct have all been studied (*Table 4.1.*). Especial interest has been directed to protein helicases, *i.e.*, unwinding enzymes[317,318] postulated to regulate G4 folding *in vivo*. Several helicases with specific G4 activity vs. dsDNA are known (*vide infra*). For example, unwinding from 5' to 3' in three ATP-dependent steps was observed with an htelo construct in presence of the yeast Pif1 helicase: i) G4 (0.77) to G-triplex (0.64), ii) G-triplex to G-hairpin (0.50), iii) G-hairpin to unfolded DNA (0.32). Importantly, unwinding of htelo by this enzyme occurred even in presence of the G4-stabilizing ligand BRACO-19[311] (*Chapter 1*). This should be taken into account for the future design of small-molecule ligands targeting G4 structures *in vivo*.

Another enzyme, the human Bloom (BLM) helicase, binds and unwinds intermolecular and intramolecular G4s using ATP[312]. BLM deficiency causes Bloom syndrome, an autosomal recessive condition related to dwarfism, immunodeficiency, and cancer[319]. G4 motifs were found to be enriched at transcription start sites and within first introns of differentially expressed mRNAs in Bloom syndrome when compared with healthy cells. This suggests that G4 structures are physiologic targets for BLM[319,320]. Upon addition of this helicase, G4 unfolding has been reported to be strongly influenced by the structural environment of a 3-quartet G4, with the enzyme requiring a 3' single-stranded overhang to be active. Unfolding of an htelo DNA took place in an ATP-dependent manner along the 3' to 5' direction and terminated with an unfolded 0.32 FRET state[321].

Unfolding by BLM in 50 mM K^+ was also observed in absence of ATP when the construct contained an internal overhang between the G4 DNA and the immobilized duplex. In parallel, a construct without the internal overhang showed that in presence of ATP the more K^+ was present and the more stable the G4 motif, the less unfolding occurred[322].

Budhathoki *et al.* further exemplified how extreme care must be taken in the design of the smFRET nucleic acid construct. The authors observed duplex unfolding with the BLM helicase in an immobilized intramolecular htelo G4 DNA with a 3' ss overhang[313]. Subsequently, they changed the polarity of the construct so that the G4 end was at the 5' end, to avoid the opening of the duplex, which would disturb the surface immobilization. Now, with a 12-T ss overhang between the G4 (at its 3') and the duplex, 3' to 5' unfolding of the G4 happened without affecting the duplex. Upon BLM addition, in absence of ATP, to the initial 0.46 and 0.62 FRET peaks for G4 at 150 mM K^+, a new FRET state was observed at 0.26 corresponding to the

BLM-bound unfolded G4. In this case, no step-wise unfolding was observed as in the case of Pif1 (*vide supra*), and instead G4 unwinding happened by a collapse of the G4 structure upon removing a few nucleotides.

The human Werner syndrome ATP-dependent helicase (WRN) showed a similar G4 unfolding activity as BLM, but not human RecQ and *E. Coli* RECQ5, which did not unwind the structure in the absence of ATP at physiological ionic strength. All these ReQ helicases share two central domains that bind and hydrolyze ATP and allow enzyme translocation on ssDNA in the 3' to 5' direction[313,314].

The reported smFRET studies on G4 DNAs summarized here confirm that cation type and concentration, annealing conditions, loop length, G4 stability, and the presence of interacting proteins and ATP, are all factors influencing G4 folding, unfolding, and dynamics. Indeed, stabilizing cations increase the folding kinetics and decrease the unfolding rates, while more stable G4s, either containing shorter loops or folded at higher monovalent cation concentrations, slow down the unfolding by a G4 helicase. At the same time, these reports are a proof-of-principle for the use of smFRET in the context of G4s, as a technique that allows detection of heterogeneities and of low-populated G4 states[137,307-309].

4.1.2 RNA construct design for smFRET

As described in *Section 4.1.*, smFRET studies allow to determine the kinetics and the folding intermediates of G4 structures. In order to elucidate the folding pathway and dynamics of the NRAS G4 by means of single-molecule FRET we designed a RNA-DNA construct following the reported strategy used successfully for G4 DNAs[302,323]. As mentioned in Section 4.1.1. the constructs are usually designed with a double-stranded overhang, each strand carrying one fluorophore. This is to avoid the costly synthesis and low yields associated with both dyes being located on the same strand[324].

Therefore, we commercially obtained a DNA oligonucleotide labeled with a biotin at the 5' end and with a sulfonated Cy3 (SCy3) internally. The sequence was chosen as complementary to the G4-forming RNA strand to allow annealing. The RNA was also obtained commercially and consisted of the 18-nt NRAS sequence with a 5' NH_2 modification and a 21-nt 3' overhang (sequence chosen from the naturally occurring mRNA, *full sequence in Materials & methods*). After 5'-labelling of the RNA with a sulfonated Cy5 (SCy5) (*Section 4.2.1.*), thermal annealing of both SCy3-DNA-biotin and SCy5-RNA oligonucleotides yielded an intramolecular NRAS G4 construct susceptible to be immobilized on a quartz surface via biotin-streptavidin binding (*Fig. 4.1.*).

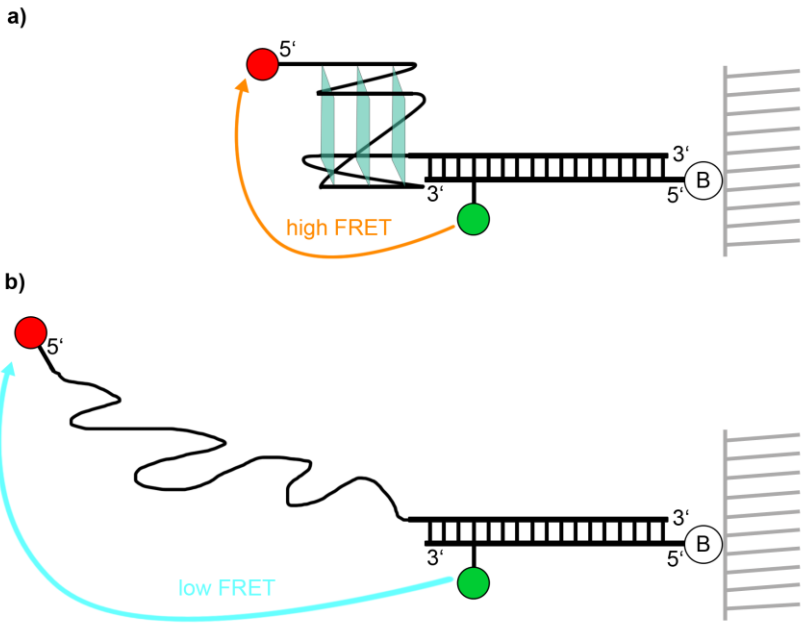


Figure 4.1. NRAS construct used in smFRET measurements. The G4-containing RNA strand is labeled with the SCy5 acceptor at the 5' and hybridized with a DNA oligomer at the 3' end, which carries the SCy3 donor as an internal modification (3-nt away from the G4) and a biotin molecule at the 5' end. Biotin allows for surface immobilization on a quartz slide via biotin-streptavidin linkage (*Section 1.3.3.*). a) Upon G4 folding, the donor and acceptor dyes are close enough (~ 3.4 nm, estimated[65,70]) for FRET to occur in high efficiency and a high FRET state is expected. b) If the G4 motif unfolds into non-structured RNA, the SCy3-SCy5 distance will become longer (~ 7.1 nm, estimated[65]) and a low FRET signal is predicted.

Cyanine dyes are chosen because of their photostability[323] and used in their sulfonated versions SCy3 and SCy5 (*Fig. 1.14.a)*. They are thus negatively charged and their water solubility is increased, at the same time expecting their interaction with the polyanionic RNA to be minimized. Upon excitation of the SCy3 donor at 532 nm, FRET at the single-molecule level can be observed if the distance between the dyes is short enough (up to ~ 10 nm, *Section 1.3.1.*). The estimated inter-dye distance between SCy5 at the G4 5' and SCy3 at 3-nt from the G4 3' is ~ 3.4 nm if the G4 is folded and ~ 7.1 nm if it unfolds (calculated from the size of G4 published structures and the distance between duplex bp)[65,70]. Therefore, high FRET is expected if the G4 is formed and stable, and lower FRET if the RNA is unstructured (*Fig. 4.1.*).

Furthermore, our setup allows to verify the presence of the SCy5 acceptor fluorophore at the single-molecule level by alternating laser excitation (ALEX), which consists in intermittently blocking the green (532 nm, direct excitation of SCy3) and the red (640 nm, direct excitation of SCy5) lasers. In this manner, artifacts or impurities emitting in the SCy5 channel can be identified and discarded (*vide infra*).

4.2 Results and discussion

4.2.1 RNA labeling with a fluorophore

5' end post-synthetic labeling[325] was achieved by NHS ester coupling of the acceptor fluorophore with a 5'-NH_2 linker on the RNA (*Fig. 4.2.*). The labeling reaction was performed in 0.1 M sodium bicarbonate buffer, pH 8.40. A tight control of the pH is crucial, as more basic conditions would result in hydrolysis of the reactive NHS ester and its inactivation (*Fig. 4.2*), while a lower pH would yield an inert protonated RNA-NH_3^+ derivative[326].

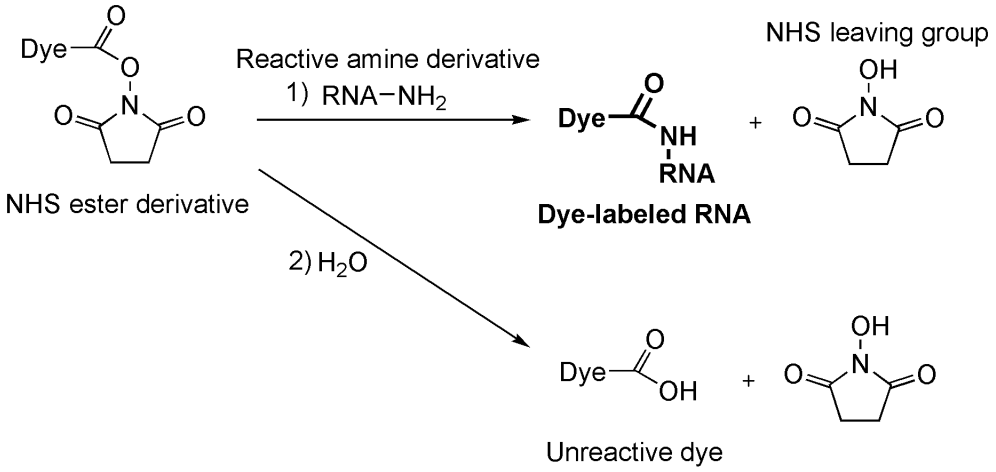


Figure 4.2. Post-synthetic RNA labeling. The labeling is performed through NHS ester coupling of the SCy5 dye to the RNA 5' end, which contains a NH_2 modification. Reaction 1) is the labeling reaction, yielding the 5'-dye-labeled RNA and the corresponding NHS leaving group. Reaction 2) shows the inactivation of the NHS ester at basic pH, by a side-reaction with water that hydrolyzes the ester to its acid form[327]. To avoid NHS ester inactivation in aqueous conditions it is important to maintain the pH in a tight range, 8.3-8.5. This is achieved with a buffered solution, e.g. 0.1 M sodium bicarbonate. At a lower pH, the amino group on the RNA would be protonated and no modification would take place[326].

Different reaction conditions were screened for the SCy5-NRAS RNA reaction, varying the following parameters:

- temperature, 11 °C or 4 °C [326,328]
- shaking, 0 or 300 rpm
- time, 24 h or overnight (~ 16 h)[326,328,329]
- excess of dye, 8 equivalents[326] or 20 equivalents

After HPLC analysis of the products, detecting the RNA and the dye simultaneously by measuring the absorbance at 260 nm for the RNA and at 646 nm for the dye, respectively, the best conditions were found to be the ones depicted in *Fig. 4.3.* The labeling reaction was performed in the minimal volume possible[326] and repeated twice by re-labeling the product of a first reaction in order to increase the final labeling efficiency. The SCy5-RNA product was detected via analytical HPLC[329] (appearing at retention time ca. 12.9-13.0 min; *Appendix,*

Figs. C.1.-C.3.) and denaturing fluorescent PAGE (unlabeled RNA and SCy5 used as controls; *Appendix, Fig. C.4.*) and subsequently purified by HLPC[329]. MALDI-MS analysis confirmed the expected 5'-SCy5-labeled NRAS RNA (*Fig. 4.4*). The yield was ~ 28 %, which is consistent with reported yields, typically between 10 and 50 %[329].

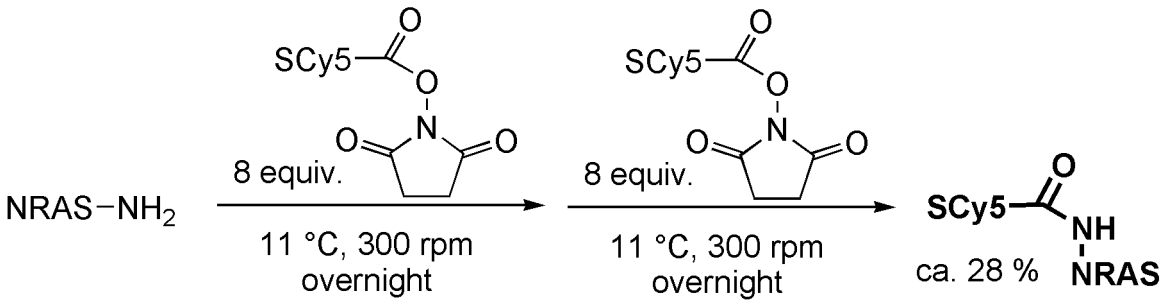


Figure 4.3. 5'-SCy5 labeling of NRAS G4. The labeling reaction was performed twice in 0.1 M sodium bicarbonate with 8 equivalents of SCy5 NHS ester by shaking overnight in a Thermomixer at 300 rpm and 11 °C. The total reaction volume was as small as possible, ~ 20-25 µL, and the SCy5-NRAS was obtained with ca. 28 % yield after HPLC purification.

After labeling was completed, the labeling efficiency of the purified RNA was calculated as the ratio of the SCy5 dye concentration, obtained from the absorbance at 646 nm, to the RNA concentration, obtained from the absorbance at 260 nm[328]. Commercially obtained SCy5-NRAS had a labeling efficiency of 99 % (for UV spectra, *Appendix, Fig. C.5.a*). Meanwhile, the in-house labeled SCy5-NRAS showed a labeling efficiency of ca. 50 % (for UV spectra, *Appendix, Fig. C.5.b*). This difference, together with the lengthy purification procedure for only a moderate yield, prompted us to use the commercial SCy5-NRAS for further smFRET experiments.

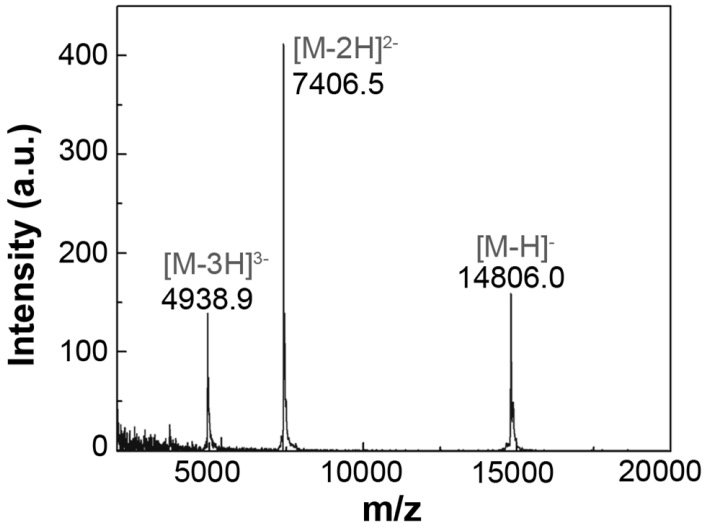


Figure 4.4. MALDI-MS spectrum of the 5'-SCy5-labeled NRAS-RNA after HPLC purification. Three clear peaks are observed: 1) 4938.9 m/z, $[M\text{-}3H]^{3-}$; 2) 7406.5 m/z, $[M\text{-}2H]^{2-}$; and 3) 14806.0, $[M\text{-}H]^{-}$, corresponding to the calculated masses of 4934.7, 7402.6 and 14806.1 m/z, respectively (expected molecular weight: 14808.23 g/mol).

4.2.2 Bulk control experiments

Quantum yields measured in our group (by Dr. Richard Börner) for the sulfonated NHS ester reactive dyes commercially obtained from Lumiprobe (Hannover, Germany), are in agreement with the literature[327]: 0.2 for SCy5 (ε_{633} = 271'000 M^{-1} cm^{-1}), and 0.1 for SCy3 (ε_{532} = 162'000 M^{-1} cm^{-1}).

To check whether the presence of the SCy5 dye destabilizes the G4, UV thermal melting spectroscopy was carried out with the 5'-labeled NRAS G4 and compared to the unlabeled 43-nt NH_2-NRAS and to the native 18nt-NRAS RNA sequence (*Table 1 and Appendix, Fig. C.7.*). Addition of the 21-nt single-stranded RNA overhang does not change the melting temperature compared to the native sequence and thus does not destabilize the NRAS G4. In contrast, the RNA with SCy5 at the 5' end has a lower melting temperature by ΔT_m = − 11 °C. Therefore, the presence of a cyanine dye destabilizes the G4 RNA. This will have to be taken into account when interpreting the smFRET results.

Table 4.2. Stability of the SCy5-NRAS G4. Melting temperature of the SCy5-NRAS RNA measured in 10 mM lithium(I) MOPS pH 7.4 supplemented with 20 mM K^+, as compared to the native NRAS G4 sequence.

M^{n+}	18-nt NRAS T_m (°C)	43-nt NH_2-NRAS T_m (°C)	43-nt SCy5-NRAS T_m (°C)
20 mM K^+	75±1	75±1	64±1

Circular dichroism confirms that the G4 forms in the SCy5-labeled and DNA-hybridized smFRET construct: the characteristic parallel G4 peaks at 243 nm and 262-264 nm are observed after 1:1 thermal annealing of the 43-nt SCy5-NRAS and the SCy3-DNA-biotin at 10 µM, in 10 mM lithium(I) MOPS buffer, pH 7.40, with 20 mM $KClO_4$ (*Appendix, Fig. C.4.*).

Several native polyacrylamide gels were run in order to establish whether the duplex is forming correctly between the NRAS RNA and the DNA carrying SCy3 and biotin. In all cases, the DNA and RNA were mixed 1:1 and heated for 5 min at 95 °C to disrupt any unwanted interactions. Different annealing conditions were tested as follows: very slow cooling to low temperature vs. fast cooling on ice; 100 mM vs. 500 mM Li^+, and 10 mM K^+ vs. no added K^+ for buffer containing 100 mM Li^+. After fluorescent detection, a red band on the gel is expected for SCy5_NRAS, a green one for SCy3-DNA-biotin, and a yellow band for the double-stranded construct containing both dyes. The results in the optimized annealing conditions, fast cooling in 100 mM Li^+ and no added K^+, show that the duplex is indeed forming as predicted, as observed by the presence of a yellow band (*Fig. 4.5., C*).

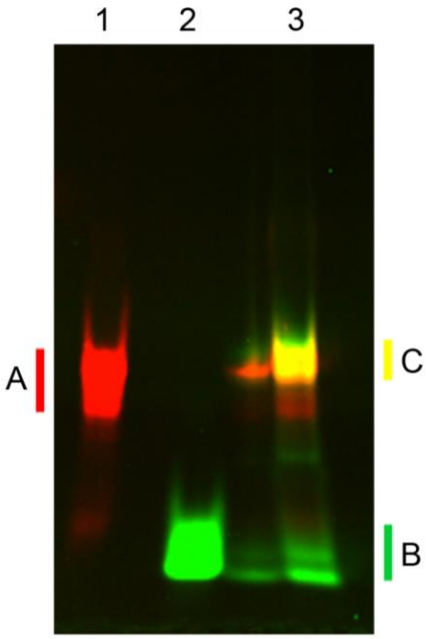


Figure 4.5. 10 % native PAGE without K^+. PAGE run at 4 °C, ca. 6h at 12 W. All samples in 100 mM Li^+, annealed at 95 °C (5 min) and quick-cooled on ice. 1) 5 pmol SCy5-NRAS, 2) 5 pmol SCy3-DNA-biotin, 3) 5 pmol SCy5-NRAS and SCy3-DNA-biotin(1:1). Band A corresponds to SCy5-NRAS RNA, band B to the SCy3-labeled DNA oligo and band C, to the hybridized RNA:DNA duplex.

4.2.3 Single-molecule FRET

All buffers for smFRET measurements are prepared containing Li^+ (500 mM) as the only cation to avoid G4-stabilizing metal ions. Indeed, Li^+ is known to stabilize G4s only very weakly[330], vs. e.g. K^+, which has a great stabilizing effect[331] (*Chapter 2*). Firstly, the background is measured, adding to the quartz slide an unlabeled RNA:DNA construct, which corresponds to the smFRET construct and lacks only the SCy5 and SCy3 dyes. The Cy3 channel appears completely clean. However, the Cy5 channel contains some fluorescent impurities, which might lead to misinterpretation of individual single-molecule fluorescent time traces (*Fig. 4.6.a*). This can easily be overcome by the use of alternating-laser excitation (ALEX)[332]. Using this technique, all molecules not showing fluorescent in both the Cy3 and Cy5 channels are discarded from further analysis, thus removing fluorescent artifacts.

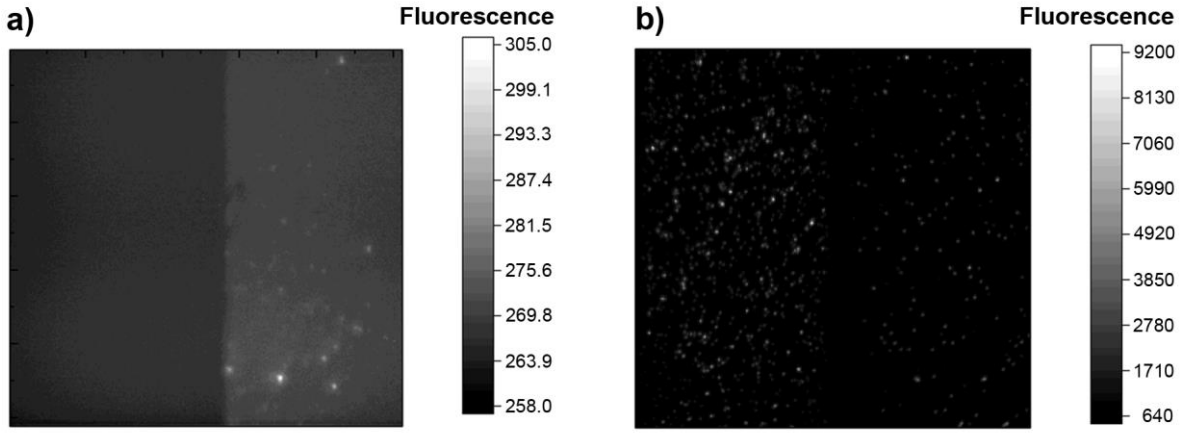


Figure 4.6. Measurement of the background fluorescence. a) 1:1 RNA:DNA unlabeled construct at 50 pM in lithium(I) MOPS, where some red-emitting fluorescent impurities are observed; b) 50 pM RNA:DNA (1:1) Cy5- and Cy3-labeled construct at 50 pM in lithium(I) MOPS containing 100 mM K^+. Some SCy3-SCy5 spots, corresponding to real bi-labeled nucleic acid molecules are detected here (40-53 % as seen by ALEX).

The density of the immobilized molecules is optimized by varying the RNA:DNA ratio from 1:1 to 1:2 and 1:0.1 (*Fig. 4.7.*). Ideally, a good density corresponds to about 150-200 molecules per observation surface[323] to ensure the correct localization of single spots. The best results are obtained with a 1:1 ratio between the two fluorophore-labeled nucleic acid strands in lithium(I) MOPS containing 100 mM K^+, which results in 40-53 % bi-labeled molecules, as controlled by ALEX. This is consistent with reported literature values of 45-85 % for the SCy3-SCy5 FRET pair[333-335]. In comparison, a 1:2 ratio yields only 11-15 % of bi-labeled molecules in the same buffer and cation conditions, while 1:0.1 results in 30-67 % bi-labeled molecules in lithium(I) MOPS.

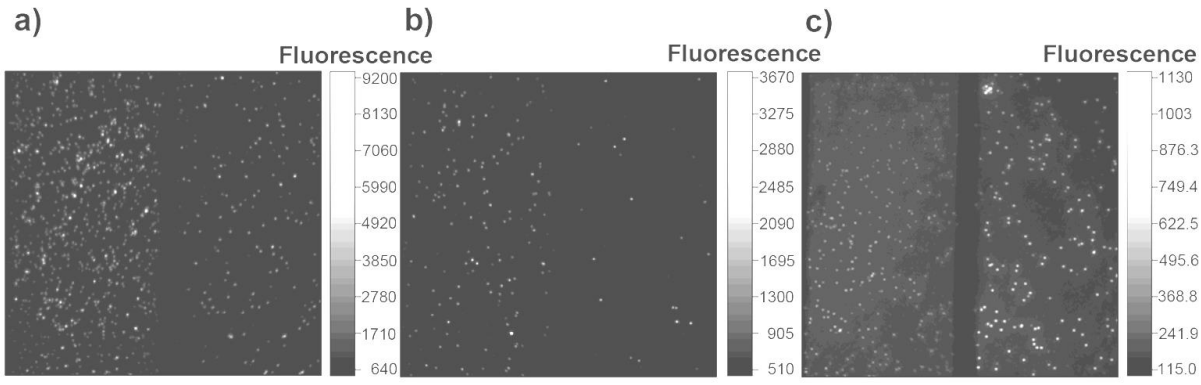


Figure 4.7. Optimization of the density of immobilized molecules. a) 1:1 RNA:DNA at 50 pM in lithium(I) MOPS containing 100 mM K^+. The density of single molecules on the slide is adequate. b) 1:2 RNA:DNA at 50 pM in Li^+ buffer containing 100 mM K^+. Too few spots appear in the SCy5 channel compared to the signal collected in the SCy3 channel. c) 1:0.1 RNA:DNA at 50 pM in Li^+ buffer. The density on the slide is acceptable here too.

smFRET measurements are carried out at room temperature (25 °C) with the labeling strategy shown in *Fig. 4.1.* and in the presence of 0-100 mM K^+. Only few molecules display changes in the FRET value or anticorrelation of the donor and acceptor intensities (donor intensity decreasing when acceptor intensity increases, or vice versa)[323] and thus almost no conformational dynamics are observed. Furthermore, no dynamics or increase in the number of anti-correlated traces are seen upon addition of K^+: 0.1, 1, 5, 25, or 100 mM.

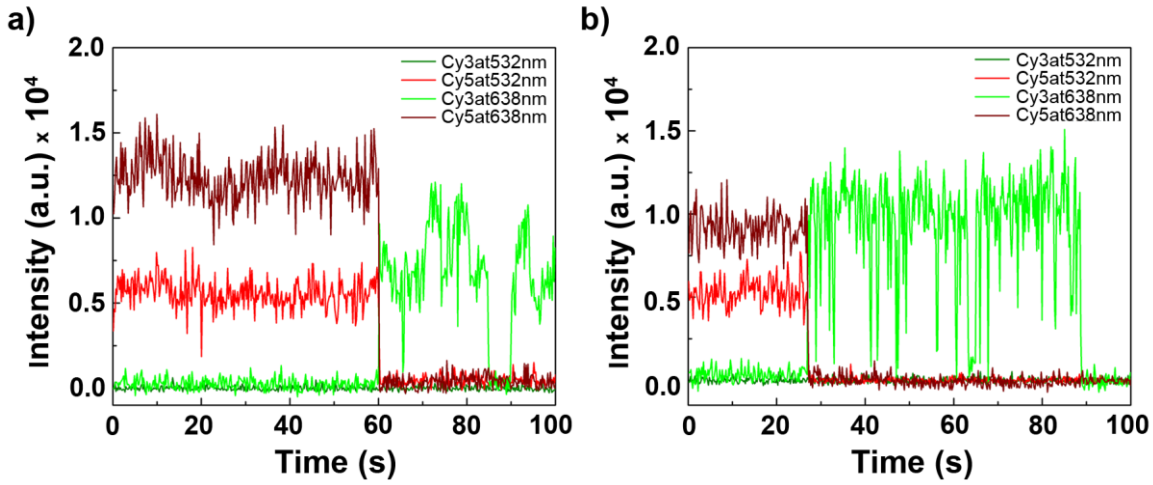


Figure 4.8. NRAS G4 time traces. Traces of fluorescence intensity vs. time obtained in lithium(I) MOPS with ALEX, at 50 pM RNA and 100 pM DNA (1:2 ratio) with 100 mM K^+. Both traces show anticorrelation (donor intensity increasing when the acceptor intensity decreases) upon SCy5 bleaching (at time ~ 60 s in a) and ~ 28 s in b)), permanent decrease in dye intensity). Considerable blinking from the SCy3 dye can also be observed (as short instances when the intensity decreases to zero, e.g. at time ~ 65 s in a) and ~ 60 s in b)).

The SCy3 and SCy5 fluorescent dyes are known to suffer from blinking and bleaching, that is, intermittent dark states and permanent loss of fluorescence, respectively[323]. Trolox and an oxygen scavenger system thus are added to the imaging buffer to minimize these effects (*Chapter 1*). However, photobleaching of SCy5 is still frequently observed in the obtained smFRET traces (7-22 % at 1:1 RNA:DNA, 9-14 % at 1:2 and 2-6 % at 10:1, *Fig. 4.8.*), while SCy3 photobleaching is not present in the time scale of the measurements (100 s). SCy3 blinking is particularly observed after SCy5 bleaching (*Fig. 4.8.*) and might be due to inter-system crossing or RNA-dye interactions leading to a change in the fluorophores' QY and therefore to changes in the fluorescence intensity (*Fig. 4.8. and Appendix, Figs. C.8. and C.9.*). SCy3 might be partially stacked onto the RNA backbone and therefore prevented from rotating freely. Measurements of the anisotropy of the dyes should be performed in the presence and absence of the nucleic acid construct in order to confirm this hypothesis.

Only the bi-labeled molecules in which the dye stoichiometry is ~ 0.4-0.6 (stoichiometry of 0.5 means both dyes are present in a 1:1 ratio) are considered to build FRET histograms. Only a high FRET state ~ 1 is observed in either absence or presence of K^+ (0.1, 1, 5 or 100 mM)

(*Fig. 4.10.*). In some traces, the fluorescence of both the donor and the acceptor fluorophores decreases to zero simultaneously for a few seconds (*Fig. 4.9.*). This might be due to dye co-quenching, which is known to happen when the fluorophores are less than 2 nm apart. This would confirm that both dyes come very close in space in the NRAS construct when the G4 is folded.

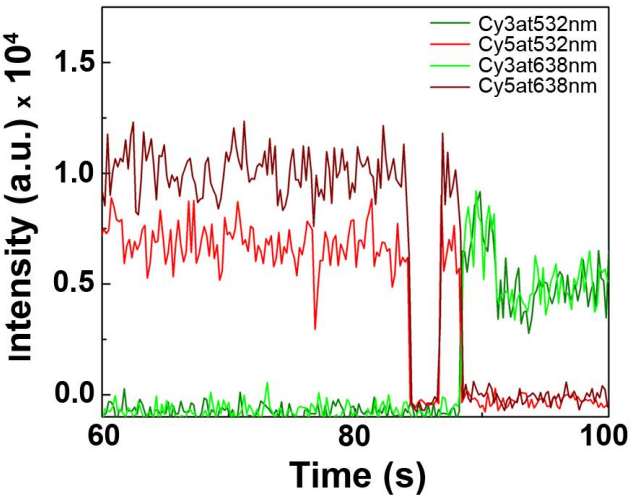


Figure 4.9. Co-quenching of the fluorophores. Time trace showing co-quenching of the SCy3 and SCy5 dyes at time ~ 85 s. The measuring conditions are the following: lithium(I) MOPS, pH 7.4, with 100 mM K^+, ALEX, 50 pM RNA and DNA (1:1).

The NRAS G-quadruplex motif is folded and stable in all cation conditions tested herein. No dynamics are observed, that is, there are no transitions between FRET states, not even in lithium(I) MOPS without added K^+. In the next section (*Section 4.2.4.*), unfolding strategies will be proposed as preferred alternatives for future smFRET experiments.

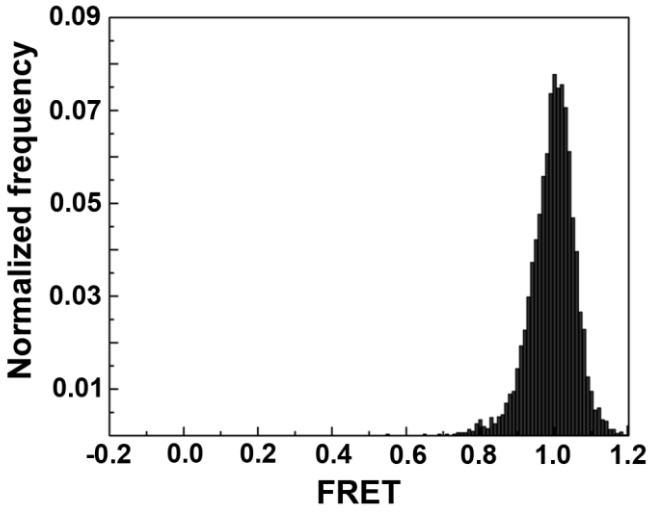


Figure 4.10. NRAS G4 smFRET histogram at 100 mM K^+. 50 pM RNA, 1:1 RNA:DNA ratio, 21 molecules, lithium(I) MOPS, pH 7.4, with 100 mM K^+.

4.2.4 Proposed unfolding strategy

As reported above, there are no dynamics observed for the NRAS G4 under the conditions tested so far. Therefore, another strategy will have to be developed in order to follow the kinetics and the folding / unfolding steps of this structure.

As a first rapid and easy option the temperature can be increased to 40-50 °C during measurements using a closed water circuit connected to the objective. However, caution should be taken that the immobilization is not perturbed and that the RNA-DNA duplex is not denatured. Another option would be to add the C-rich RNA or DNA strand complementary to the G4-forming region in order to force the unfolding of the G4. Similarly, the DNA oligo might also be elongated towards the G4, with one or more C-rich tracts complementary to the RNA G-runs.

A very attractive solution would be to add a naturally-occurring G4-re2lated protein: the G4 resolvase (G4R1), also called RHAU (RNA Helicase associated with AU-rich element), an ATP-dependent enzyme known to have a specific helicase activity upon G4 structures[336]. It shows a high affinity and specificity for RNA G4s with a K_d in the pM range[318] and is believed to have a regulatory G4-related function *in vivo*[337]. An elegant strategy would therefore be to perform smFRET competition experiments together with the G4 resolvase (*Fig. 4.12.*) and G4-stabilizing ligands. This would especially be relevant from the point of view of *in vivo* G4 folding and function.

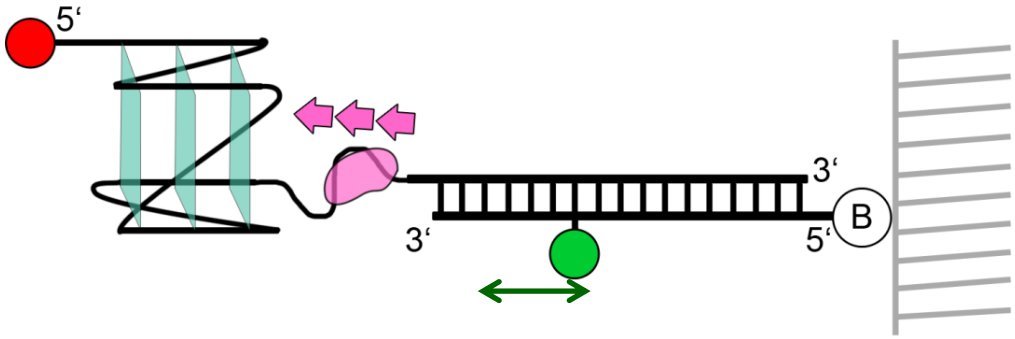


Figure 4.11. Proposed strategy for smFRET unfolding experiments with G4R1. The RNA:DNA construct is mainly the same from *Fig. 4.1* and *Section 4.2.3.*, to which a single-stranded linker between the duplex and the G4 has been added to allow the anchoring of the helicase. The D-A distance would have to be optimized by moving SCy3 along the DNA oligo in order to avoid co-quenching of the dyes.

G4 resolvase

G4R1 is the product of human gene DHX36[338] and was first identified and characterized by Harrington *et al.* in 1997 [339]. It requires a 3’ ssDNA overhang to unwind G4s[340] and has a binding preference for parallel motifs[307] unfolding both G4 DNAs and RNAs.

G4R1 is highly conserved in vertebrates and is critical for embryonic development and hematopoiesis. Its unwinding of G4 DNA is vital for spermatogenesis and for c-kit expression; a gene involved in this process, which is also an oncogene in some types of cancer[341]. G4 resolvase has also been shown to unwind an intramolecular G4 RNA at the 5' end of the human telomerase[342] thus stimulating telomerase-mediated telomere m2aintenance[343]. Finally, this enzyme is the major source of tetramolecular G4 DNA and RNA resolving activity in human HeLa cell lysates (HeLa cells are rapidly growing and contain a high concentration of G4R1)[336,338].

The enzyme has a higher affinity for G4 RNA with apparent K_d values for tetramolecular motifs of 39±6 pM for RNA, and 77±6 pM for DNA[336]. It loads on the 3′-end of G4 structures and translocates in the 3′ to 5′ direction[340]. ATP-dependent G4 unfolding is achieved at 50-60 % efficiency under physiological salt and pH conditions *in vitro*[313].

There is an inverse correlation between the thermodynamic stability of the G4 substrates and the rate of unwinding by G4R1[344] as evidenced by parallel, tetramolecular G4 DNAs $[TTA\underline{G}_nA_{15}]_4$ with n= 5-8 and by the same motifs with added G4-stabilizing ligands. However, ATP consumption is independent of G4 thermal stability. The same authors proposed a non-processive and non-translocating G4 unfolding mechanism[344]:

- G4R1 binds to the nucleic acid 3' tail in an ATP-independent manner;
- upon ATP hydrolysis the helicase destabilizes some of the G-tetrads and subsequently dissociates;
- depending on the G4 thermodynamic stability the partially unbound substrate either re-anneals into a G4 or unwinds into a single strand.

This mechanism is similar to that of DEAD-box helicases acting on double-stranded nucleic acids, which is consistent with the structural conservation in the core and N/C terminal domains between both enzyme types. The core region of ca. 470 amino acids is responsible for ATP binding, ATP hydrolysis, and helicase activity[344]. It contacts the substrate via the sugar-phosphate backbone and therefore does not induce any structural G4 selectivity. In contrast, the N-terminus of G4R1 contains an element, called RHAU-specific motif (RSM), which is responsible for specific G4 recognition and binding[344] and mediates the interaction with a guanine tetrad plane with no effect of the RNA 2'-OH[344].

G4R1 has been shown to suppress expression of the transcription factors PITX1[346] and Yingyang[337] *in cellulo*. Therefore, the unwinding by G4R1 may play a fine-tuning role in controlling gene expression *in vivo*[344].

4.3 Conclusions

In this Chapter 4, the aim was to establish a working system to observe the NRAS G4 by smFRET, in order to elucidate its folding steps and dynamics. No smFRET studies on G4 RNA have been reported so far. Therefore, the construct design is based on the literature examples for G4 DNAs.

The successful 5' labeling of the NRAS G4 RNA with SCy5 is described, needing a re-labeling at 11 °C with 8 equivalents of dye. The fluorescent product obtained in ~ 28 % yield is confirmed by HPLC and fluorescent native PAGE gels and detected by MALDI-MS after HPLC purification. However, due to the sub-optimal labeling efficiency (ca. 50 %) SCy5-NRAS is finally obtained commercially and shows 99 % of the RNA molecules containing the dye.

As a first step towards kinetic and dynamic studies of the NRAS G4 via smFRET, the correct immobilization of the 1:1 RNA:DNA construct is shown with an optimum density of molecules on the observation surface. The number of bi-labeled SCy3-SCy5 molecules detected by ALEX in lithium(I) MOPS containing 100 mM K^+ represents 40-53 % of the total. The fact that half of the molecules contain only the donor fluorophore is thought to be due to acceptor pre-bleaching (that is, permanent suppression of SCy5 fluorescence prior to the measurements)[334] although it may also be due to partial RNA:DNA annealing.

Anticorrelation is observed upon SCy5 photobleaching in 2-22 % of the smFRET traces. Taking into account only bi-labeled molecules with 1:1 dye ratio, a high FRET state ~ 1 is observed in lithium(I) MOPS containing 100 mM K^+. This state, seen also in Li^+ only, likely corresponds to the folded, parallel, NRAS G4 seen in CD spectra under the same conditions (*vide supra*). No molecular dynamics, not even at lower K^+ concentration or without any K^+ present, are observed so far presumably due to high G4 stability.

Therefore, an outlook strategy has been proposed, in which unfolding of the G4 is forced by addition of a G4 resolvase. This approach, which would allow competition experiments in presence of G4-stabilizing ligands, would give an insight into the G4 RNA unfolding kinetics and would be most relevant for the understanding of RNA G4 behavior *in vivo.*

Chapter 5

Experimental section

5.1 Materials and instruments

5.1.1 Chemicals

Sample preparation

DNA templates and nucleotide triphosphates (NTPs) for *in vitro* transcription of RNA were obtained from Microsynth (Balgach, Switzerland) and GE Healtchcare (Glattbrug, Switzerland), respectively. The T7 polymerase was homemade[347]. All chemicals for transcription solutions were at least puriss. p.a. grade (or molecular biology grade) and obtained from Sigma-Aldrich (Buchs, Switzerland), Merck (Zug, Switzerland), or Acros Organics (Geel, Belgium).

All samples, buffers, and reactions were prepared in deionized RNAse-free water obtained from a TKA GenPure water purification system from TKA Wasseraufbereitungssysteme (Niederelbert, Germany), subsequently filtered and autoclaved. In addition, solutions were autoclaved (where possible) and/or sterile filtered before use with 0.2 µm Filtropur syringe filters from Sarstedt AG (Nümbrecht, Germany) or 0.22 µm Steritop bottle-top filter units from Merck Millipore (Billerica, MA, USA) to prevent RNA degradation.

Spectroscopy

Metal salts, 3-(*N*-morpholino)propanesulfonic acid (MOPS), Tris-HCl, boric acid, *N*,*N*,*N'*,*N'*-ethylenediaminetetraacetic acid (EDTA), 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), tetrabutylammonium (TBA) hydroxide and 5,10,15,20-tetrakis(1-methyl-4-pyridinio)porphyrin tetra(*p*-toluenesulfonate) (TMPyP4) were at least puriss. p.a. grade (or molecular biology grade) and obtained from Sigma-Aldrich (Buchs, Switzerland), Merck (Zug, Switzerland), or Acros Organics (Geel, Belgium).

Deuterated solvents for NMR spectroscopy*, i.e.* D_2O and d_3-acetonitrile, were from Armar Chemicals (Döttingen, Switzerland).

Electrophoresis

AccuGel (29:1 acrylamide:bisacylamide) for polyacrylamide gel solutions was obtained from National Diagnostics (Hessle, UK). SybrGold for staining of the gels was from Life Technologies (Zug, Switzerland).

smFRET

The SCy5 NHS ester for RNA 5' labeling was obtained from Lumiprobe (Hannover, Germany).

5.1.2 Buffers

All buffers were prepared in deionized, autoclaved, RNAse-free water. This same type of water was always used to dissolve DNA or RNA samples, unless otherwise stated.

Transcription buffer: 40 mM Tris-HCl (pH 7.5), 40 mM dithiothreitol (DTT), 2 mM spermidine, 0.01 % Triton X-100.

TBE buffer for denaturing PAGE and electroelution: diluted to 1 × TBE from commercial 10 × TBE from National Diagnostics (Hessle, UK).

Urea loading buffer for denaturing PAGE: 11.7 M urea, 40 mM Tris-HCl (pH 7.5), 0.1 % xylene cyanol (XC), 0.1 % bromophenol blue (BB), 230 mM sucrose, 0.8 mM EDTA pH 8.0.

Native PAGE buffers: 1 × TBE, pH 8.3, diluted from 10 × TBE without cations, prepared from Tris(hydroxymethyl)aminomethane (Tris), boric acid and EDTA. The final 1 × TBE was supplemented with 20 mM KCl, 20 mM $SrCl_2$ or 100 mM NH_4OAc.

Native buffer for fluorescent PAGE: 66 mM HEPES, 34 mM Tris, 50 mM MOPS and 100 mM LiCl. Final pH adjusted to 7.4.

Lithium(I) MOPS: 10 mM MOPS + 2 M LiOH until pH 7.40. The final lithium(I) concentration is 10 mM.

Sodium(I) MOPS: 1 mM MOPS + 5 M NaOH until pH 7.40. The final sodium(I) concentration is 1 mM.

TBA MOPS: 10 mM MOPS + 40 % tetrabutylammonium hydroxide (TBAOH) until pH 7.40. The final TBA concentration is 10 mM.

T50 buffer: 50 mM Tris-HCl (pH 7.5), 50 mM LiCl.

Standard buffer for smFRET: 50 mM lithium(I) MOPS, 450 mM LiCl, pH 7.4.

Oxygen scavenger solution (OSS): 100 × stock, 2170 U/mL catalase and 165 U/mL glucose oxidase dissolved in 80 µL T50 buffer.

Imaging buffer for smFRET: 1mM Trolox, 1 × OSS, 1 % (w/v) D-glucose, dissolved in standard buffer, pH 7.4.

Potassium buffer for NMR: 100 mM potassium(I) phosphate buffer stock, pH 7.4, prepared with K_2HPO_4 and KH_2PO_4 at ratio 80.2:19.8.

5.1.3 Oligonucleotides

The G4 RNA and DNA oligonucleotides were purchased HPLC-purified and desalted from Microsynth (Balgach, Switzerland) or IBA (Göttingen, Germany). The NRAS RNA sequence was also transcribed *in vitro* using homemade T7 polymerase[347], according to a published procedure[341]. 0.7 µM DNA template, 30 mM $MgCl_2$, 5 mM of each NTP (ATP, CTP, GTP, and UTP), and 20 µL/mL T7 RNA polymerase were mixed in 1 × transcription buffer. The reaction was allowed to proceed for 2.5 h at 37 °C whilst shaking at 300 rpm. Purification and subsequent desalting were undertaken via 20 % polyacrylamide gel electrophoresis (PAGE), ultrafiltration with Centricon Centrifugal Filter Devices (3000 MWCO) from Amicon (Merck Millipore, Billerica, MA, USA) and size-exclusion chromatography with NAP-10 columns from GE Healthcare (Glattbrugg, Switzerland). Between denaturing PAGE and ultrafiltration, the RNA was extracted from the gels by electroelution[348]. The average yield of an NRAS transcription was 8±4 nmol RNA / mL transcription mix and the final isolated NRAS RNA was analyzed by MALDI-MS in the negative-ion mode (calcd. 6216.8; meas. 6216.7 m/z).

As DNA templates for NRAS in vitro transcription, two commercial DNA sequences were used. They anneal together in solution, forming the TATA box double-stranded sequence (marked in grey) for attachment of the T7 polymerase. “ts” stands for top strand, and “ot”, for original template, which is the sequence transcribed into the targeted RNA. At the 5’ end of the ot strand, the two first bases are methylated into 2’-OMe-RNA (marked with *), which favors a homogeneous 3’ end in the transcript RNA[349].

NRAS01_ts: 5'-GAAATTAATACGACTCACTATAGGGAGGGGCGGGTCTGGG-3'

NRAS01_ot: 3'-CTTTAATTATGCTGAGTGATATCCCTCCCCGCCCAGAC**-5'

Ethanol precipitation was used to isolate the RNA from a solution or mixture, e.g. after electroelution, by adding 3 × sample volume of ice-cold ethanol and 1/20 × sample volume of 5 M NaCl or LiCl. After centrifugation (30 min at 13’200 rpm and 4 °C) the supernatant was discarded and the precipitated RNA was dried in a Concentrator 5301 from Eppendorf (Vadaux-Eppendorf, Schönenbuch, Switzerland) and re-dissolved in water before ultrafiltration and size-exclusion chromatography (*vide supra*).

The concentration of all RNAs dissolved in water was determined by UV absorption at 260 nm using the extinction coefficients at 260 nm (ε_{260}), as supplied from Microsynth (Balgach, Switzerland) (*Table 5.2.*).

Table 5.3. Nucleic acid sequences. Obtained from Mycrosynth (Balgach, Switzerland).

Name	Nucleic acid	Sequence (5'-3')	ε_{260} ($mM^{-1}cm^{-1}$)	Chapter
dNRAS	DNA	GGGAGGGGCGGGTCTGGG	179.7	2, 3
htelo	DNA	TTAGGGTTAGGGTTAGGGTTAGGG	245.2	2, 3
NRAS(18)	RNA	GGGAGGGGCGGGUCUGGG	202.3	2, 3, 4
NRAS20	RNA	UGGGAGGGGCGGGUCUGGGU	221.3	3
NRAS21	RNA	UUAGGGAGGGGCGGGUCUGGG	236.7	3
NRAS22	RNA	GUGGGAGGGGCGGGUCUGGGUG	245.5	3
TERRA(24)	RNA	UUAGGGUUAGGGUUAGGGUUAGGG	281.2	2, 3
TERRA21	RNA	GGGUUAGGGUUAGGGUUAGGG	245.4	3

For smFRET measurements, fluorophore-labeled RNA or DNA were obtained from IBA (Göttingen, Germany). Both the amino-derivative and the SCy5-labeled RNA contain a 6-carbon linker at the 5' end (C6):

NH_2-NRAS

5'-C_6NH_2-GGGAGGGGCGGGUCUGGGUGCAGCCUGCCGCAUGACUCGUGGU-3'

SCy5-NRAS

5'-SCy5C_6-GGGAGGGGCGGGUCUGGGUGCAGCCUGCCGCAUGACUCGUGGU-3'

SCy3-DNA-biotin

3'-ACGT(Cy3)CGGACGGCGTACTGAGCACCA-biotin-5'

See Chapter 4 for the labeling scheme chosen for smFRET measurements. The following extinction coefficients, as supplied by IBA, were used: 407.1 $mM^{-1}cm^{-1}$ for the 43-nt NRAS RNA at 260 nm (ε_{260}), and 271 $mM^{-1}cm^{-1}$ for the SCy5 dye at 646 nm (ε_{646}).

5.1.4 Instruments

Sample preparation

The electroelution apparatus *Biotrap* was from Schleicher & Schuell (Dassel, Germany) and was used with BT1 and BT2 membranes from Whatman (London, UK).

Nucleic acid samples were vacuum-dried with a Concentrator 5301 from Eppendorf (Vadaux-Eppendorf, Schönenbuch, Switzerland) or a freeze dryer Alpha 2-4 LDplus from Christ (Osterode am Harz, Germany).

The centrifuges Sorvall RC 6 Plus from Thermo Scientific (Waltham, MA, USA) with rotor SA-600, Eppendorf 5804R with rotor A-4-44, Eppendorf 5415R with rotor F-45-24-11 (both Vadaux-Eppendorf, Schönenbuch, Switzerland) and GmCLab from Gilson (Villiers Le Bel, France) were used. The small vortexing device PV-1 from Faust (Schaffhausen, Switzerland) was also utilized.

For heating and mixing of samples a Thermomixer Compact or Comfort from Vadaux-Eppendorf (Schönenbuch, Switzerland) were used.

pH measurements were carried out using a Metrohm (Herisau, Switzerland) 605 pH-meter and a minitrode glass electrode from Hamilton (Bonaduz, Switzerland). Calibration was performed with solutions of certified pH (pH = 4, 7, 9) from Metrohm.

To autoclave both glassware and buffer solutions, an autoclave Systec VX-100 from Systec (Linden, Germany) and a Varioklav Steam Sterilizer from Sterico (Wangen, Switzerland) were used.

Spectroscopy and spectrometry

NMR spectra were recorded on a Bruker Avance 600 MHz spectrometer with a 5 mm CRYO TCI inverse triple-resonance probehead with z-gradient coil or on a Bruker Avance 700 MHz spectrometer with a 5 mm CRYO TXI inverse triple-resonance probe-head with z-gradient coil (both: Bruker Biospin AG, Fällanden, Switzerland) at the NMR facility of the Department of Chemistry at the University of Zurich.

UV-absorption measurements were performed on Cary 100 UV/Vis, Cary 500 scan UV/Vis-NIR (both: Varian, Palo Alto, CA, USA) and Lambda 850 UV/Vis (PerkinEler, Schwerzenbach, Switzerland) spectrophotometers and on a NanoDrop microvolume UV/vis. spectrophotometer from Thermo Fisher Scientific (Waltham, MA, USA). The cuvettes were purchased from Hellma (Müllheim, Germany) or Starna (Hainault, UK).

Circular dichroism (CD) was recorded with a Jasco 810 spectropolarimeter from Jasco (Easton, MD, USA) equipped with a peltier thermostat.

Matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS) spectra were recorded at the Functional Genomics Center Zurich (FGCZ) by Dr. Serge Chesnov.

Dynamic light scattering (DLS) was measured with a DynaPro Titan spectrometer connected to a temperature controlled microsampler, both from Wyatt Technology Corporation (Santa Barbara, CA, USA).

Electrophoresis

The gel electrophoresis apparatuses were homemade (chemical workshop, UZH) and the power supplies, from Thermo Fisher Scientific (Waltham, MA, USA).

Imaging of fluorescent gels or SybrGold-stained gels was performed on a Typhoon FLA 9500 imager from GE Healthcare (Glattbrug, Switzerland). Pictures of analytical PAGE gels were taken by UV shadowing with a Bio-Vision 3026 WL/LC/26Mx gel imager system from Vilber Lourmat (Eberhardzell, Germany).

HPLC

HPLC analyses were performed on a Ultimate 3000 system from Dionex (Olten, Switzerland) equipped with a diode array and fluorescence detector, or on a LaChrom Elite system from Hitachi (Tokyo, Japan). Analytical HPLC was undertaken with a Waters (Baden-Dättwil, Switzerland) XTerra RP8 5 µm column. Preparative HPLC was carried out on a Varian (Palo Alto, CA, USA) ProStar system with two solvent delivery system units model 215 and a UV/Vis detector model 320, equipped with a Zorbax Prep HT 300SB-C18 5 µm column from Agilent (Basel, Switzerland) or a Waters XTerra Prep RP8 5 µm column.

smFRET

Single-molecule FRET measurements were performed with a total internal reflection fluorescence microscope (TIRF) system using a home-built setup according to Zhao and Rueda[224]. The setup mainly includes an IX71 microscope from Olympus (Volketswil, Switzerland), a green laser at 532 nm and a red laser at 640 nm, both from CrystaLaser (Reno, NV, USA). The output intensity after excitation of the fluorophores was attenuated using neutral density filters from Laser2000 GmbH (Wessling, Germany). Dichroic mirrors from AHF AG (Tübingen, Germany) were used for spatial separation of the fluorophore emissions. Signals were recorded with an EMCCD camera DU 897 from Andor (Belfast, UK).

Data treatment

NMR spectra were processed using TopSpin 3.0 software (Bruker Biospin AG, Switzerland).

Fluorescent or SybrGold-stained gels processing was performed with Image Quant TL from GE Healthcare (Glattbrug, Switzerland).

MOLMOL was used to display DNA and RNA structures with data obtained from the PDB.

smFRET data analysis was performed using the home-written MATLAB-based (MathWorks, Nattick, MA, USA) software called MASH (Multifunctional Analysis Software for Heterogeneous smFRET data)[350].

5.2 Methods

5.2.1 UV thermal melting experiments

UV melting curves were recorded by measuring the absorbance at 295 nm – the wavelength at which ΔA is maximal for G4[351] – as a function of temperature. Zero-absorption at 405 nm was used as a control to discard any artifacts from the measurements[352]. RNA samples were prepared with a final concentration of 4 µM in 10 mM lithium(I) MOPS buffer, pH 7.4, containing 20 mM MNO_3 (M = Li^+, Na^+, K^+, Rb^+, Cs^+ or NH_4^+) or 0.1 mM $M'(NO_3)_2$ (M' = Mg^{2+}, Ca^{2+}, Sr^{2+}, Ba^{2+}). Metal ions were pipetted from a 1 M stock solution of the corresponding nitrate salt.

Controls in either 1 mM sodium(I) MOPS, pH 7.4, 10 mM lithium(I) cacodylate, pH 7.4, or water, pH 6.4, were measured at 4 µM RNA in absence of added cations, or with 20 mM Li^+ or K^+. Finally, measurements at 1-50 µM RNA in either 20 mM K^+ or 0.1 mM Sr^{2+} were recorded.

The melting experiments were performed in 10 mm path length quartz cuvettes with 150 µL sample solution, covered with 400 µL paraffin oil to prevent evaporation. The samples were degassed and pre-heated at 90 °C, waiting 20 min for temperature stabilization before starting the measurements. Temperature ramps were recorded between 90 and 10 °C. A ramp rate of 0.25-0.35 °C/min was used and data points were collected every 1 °C. Melting temperatures were calculated by the baseline method as previously described[352] (*Fig. 5.1.*), except for 25-80 mM of Sr^{2+} and K^+, where no baseline at high temperature was observed anymore. In these cases, the melting temperature (T_m) was obtained as the maximum of the first derivative of the absorbance signal[352]. The results of three independent replicates were averaged in all cases as a weighted mean (so that more precise results have more importance), and errors were calculated as weighted standard deviations.

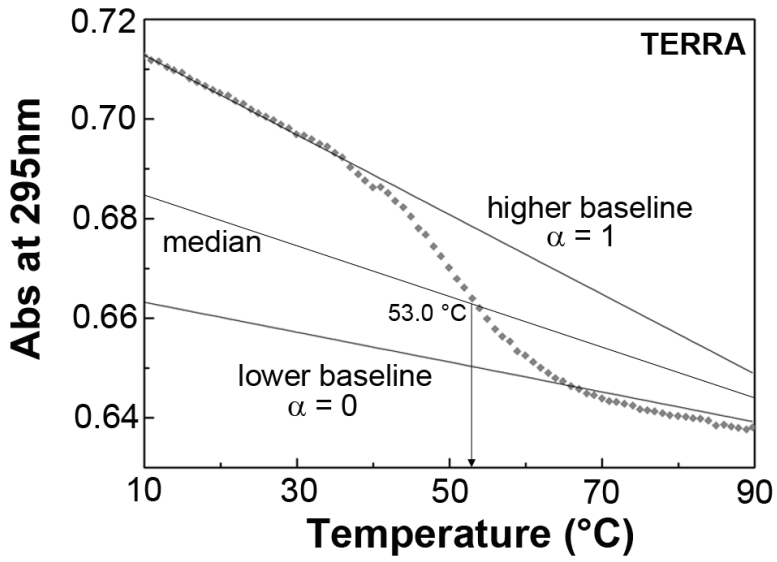


Figure 5.1. T_m determination. To obtain the melting temperature from a G4 RNA melting profile recorded at 295 nm, two baselines are fit to the low temperature and high temperature ranges, respectively. The lower baseline corresponds to the dissociated form (folded fraction $\alpha = 0$) and the higher baseline corresponds to the folded conformation (folded fraction $\alpha = 1$). A median line is subsequently drawn from the two baselines (by observing the Y-incercept values of the two baselines at two different temperatures; the median will be then drawn in the middle), and its intersection of with the melting curve yields the melting temperature[352]. In this example, the sample was 24-nt TERRA (10 µM) in 10 mM lithium(I) MOPS, pH 7.4, supplemented with 20 mM Na^+.

The thermodynamic parameters ΔG_{VH}, ΔH_{VH} and ΔS_{VH} were determined from the obtained melting curves by assuming an intramolecular G4 formation in a two-state equilibrium. The absorbance vs. temperature plot was converted into a fraction folded (α) vs. temperature representation, using Equation 5.1.:

$$\alpha_T = (L0_T - A_T)/(L0_T - L1_T) \quad (5.1.)$$

where $L0_T$ and $L1_T$ correspond to the lower and higher baseline values, respectively. The affinity constant (K_a) was then obtained from the folded fraction according to the following equation:

$$K_a = \alpha_T/(1 - \alpha_T) \quad (5.2.)$$

Subsequently, the natural logarithm of K_a ($\ln(K_a)$) was plotted as a function of the reciprocal of the temperature (1/T in K^{-1}), restricting the analysis to the temperature range for which $0.03 < \alpha < 0.97$[346]. The obtained plot, called van't Hoff representation[189], gives a straight line with a slope of $-\Delta H°/R$ and a Y-axis intercept of $\Delta S°/R$. Finally, $\Delta G°$ was calculated at 37 °C following:

$$\Delta G^{\circ} = \Delta H° - T \cdot \Delta S° \quad (5.3.)$$

5.2.2 Circular dichroism spectroscopy (CD)2

A 1 mm path length quartz cuvette was used with an RNA concentration of 10 µM in either pure water or 1 mM sodium(I) MOPS buffer, pH 7.4, containing 20 mM MNO_3 (M = Li^+, Na^+, K^+, Rb^+, Cs^+, NH_4^+) or 0.1 mM $M'(NO_3)_2$ (M' = Mg^{2+}, Ca^{2+}, Sr^{2+}, Ba^{2+}). Before measurements, the samples were annealed by heating at 95 °C for 5 min and allowed to slowly cool down to room temperature. Continuous scans were performed between 230 and 320 nm with a scanning speed of 100 nm/min, a data pitch of 0.5 nm, a response of 0.5 s, and a band width of 1 nm. Five scans were averaged, a buffer-only blank was subtracted from each spectrum, and data were zero-corrected at 320 nm. Metal ions were pipetted from a 1 M stock solution of the corresponding nitrate salt. Independent triplicates were performed of each measurement in order to ensure reproducibility. Titration curves plotted as $\Delta\varepsilon$ at 265 nm vs. M^+ concentration correspond to the arithmetic mean of three triplicates, showing the corresponding standard deviations as error bars.

In all graphs, circular dichroism is represented as molar circular dichroism ($\Delta\varepsilon$), calculated as:

$$\Delta\varepsilon = \varepsilon LCP - \varepsilon RCP = \frac{\Delta A}{(c \cdot l)}[M^{-1} \cdot cm^{-1}] \quad (5.4.)$$

where εLCP and εLCP are the molar extinction coefficients for left- and right-circularly polarized light, respectively; ΔA is the differential absorbance of left- and right-circularly

polarized light; *l* is the path length in centimeters; and *c* is the molar concentration in mol/L. The final plotted data is concentration independent[353].

For the CD titrations with K^+ or Na^+, 15 µM of RNA were dissolved in either water or 10 mM lithium(I) MOPS buffer, pH 7.4, and the corresponding cation (K^+, Na^+) was added from a nitrate stock solution, to obtain final concentrations ranging from 0-300 mM. For Sr^{2+} titrations, 10 mM tetrabutylammonium (TBA) MOPS buffer, pH 7.4., was used together with 0-100 mM Sr^{2+}. All other measurement and analyzing parameters were the same as mentioned above.

5.2.3 Thermal difference spectra (TDS)

UV spectra between 230-320 nm were recorded at 90 °C and 20 °C in 10 mm path length quartz cuvettes with 200 µL sample solution. The DNA or RNA concentration was 4-15 µM in 10 mM lithium(I) MOPS buffer, pH 7.4, containing 20 mM K^+. Subsequently, the spectrum at low temperature (20 °C) was subtracted from the one at high temperature (90 °C) for each nucleic acid sequence. The absorption difference (ΔA) was represented vs. the wavelength, and normalized by dividing all points by the highest positive value, such that the resulting normalized spectra are independent of DNA or RNA concentration.

5.2.4 Native polyacrylamide gel electrophoresis (PAGE)

20 % polyacrylamide gels were prepared containing 20 mM KCl, 100 mM NH_4OAc or 20 mM Sr_2Cl, and allowed to polymerize at 4 °C overnight. Gels were pre-run with 1 × TBE buffer, containing 20 mM KCl, 100 mM NH_4OAc or 20mM Sr_2Cl, at 12 W and 4 °C for 30 min before sample loading.

All nucleic acid sequences were first re-purified by PAGE followed by ultrafiltration and desalting with size-exclusion chromatography. Subsequently, the DNA and RNA samples in water were annealed at a concentration of 2.5 µM, with 20 mM KCl, 100 mM NH_4OAc or 20 mM Sr_2Cl, by heating at 95 °C for 5 min and cooling quickly on ice. After annealing, they were kept at 4 °C 1-2 h or overnight, after which 10 µL of 60 % glycerol (1:1 v/v) were added, mixing well. The total 20 µL of each RNA sample were loaded in the gel, corresponding to 25 pmol/well.

A commercial ds DNA ladder, O'RangeRuler 5 bp from Life Technologies (Zug, Switzerland), was used as a size marker and loaded at 1 µL/well in every gel.

The gels were run at 5 W and 4 °C for 7 h and subsequently stained for 10 to 15 min with 10 µL SybrGold in 100 mL 1 × TBE with the corresponding cation (20 mM KCl, 100 mM NH_4OAc or 20 mM Sr_2Cl).

As part of a short scientific stay at Imperial College London, in collaboration with the group of Prof. Ramon Vilar, different native PAGE gels were run. 20 % polyacrylamide gels were

prepared containing either no salt or 20 mM KCl, using a Mini-protean gel apparatus (Bio-Rad, Munich, Germany), and allowed to polymerize at 4 °C overnight. Gels were pre-run with 1 × TBE buffer containing either no salt or 20 mM KCl, at 90 V and 4 °C for 30 min. RNA samples were annealed in water at a concentration of 5 µM by heating at 95 °C for 5 min and cooling quickly on ice. Subsequently, 20 mM KCl for K^+ gels or water for gels without salt were added, together with 1:1 v/v glycerol. 10 µL of each sample were loaded into the gel wells, corresponding to 10 pmol/well. Apart from the 5 bp DNA ladder described above, a ssDNA DNA ladder was also used as a second size marker, consisting of a mixture of 15-T, 30-T and 75-T oligonucleotides (where T is thymine).

The gels were run at 90 V and 4 °C for 3-3.5 h and subsequently stained for 10 min with 10 µL SybrGold in 50 mL 1 × TBE with 20 mM KCl. Gel images were obtained with an Ettan DIGE Imager from GE Healthcare (Pollards Wood, UK).

5.2.5 Fluorescent native PAGE

10 % polyacrylamide gels were prepared, containing 66 mM HEPES, 34 mM Tris, 50 mM MOPS and 100 mM LiCl, and left to polymerize at 4 °C overnight. Gels were pre-run at 15 W for 20 min before loading the RNA samples.

The RNA samples dissolved in RNAse-free water were annealed at 500 nM with 100 mM $LiNO_3$ by heating at 95 °C for 5 min and cooling quickly on ice. 10 µL of 60 % glycerol (1:1 v/v) were added, mixing well. The total 20 µL of each RNA sample were loaded in the gel, corresponding to 5 pmol/well.

The gels were run at 12 W and 4 °C for 6 h and subsequently RNA bands were visualized by UV shadowing and by detecting the SCy3 and SCy5 fluorescence.

5.2.6 Denaturing polyacrylamide gel electrophoresis (PAGE)

20 % polyacrylamide gels containing 7 M urea were prepared and allowed to polymerize at r.t. After one hour, gels were pre-run with 1 × TBE buffer at 15-16 W for 30-60 min before loading the samples.

The DNA and RNA samples dissolved in RNAse-free water were annealed at 5 µM by heating at 95 °C for 5 min and cooling quickly on ice. Afterwards, 5 µL of urea loading buffer (1:1 v/v, *vide supra*) were added to 5 µL of annealed RNA, mixed well, and the samples heated at 90 °C for 2 min. 10 µL were loaded in the gel yielding a final RNA amount of 25 pmol/well.

The gels were run at 12-13 W and r.t. for 1.5 to 2 h and subsequently stained for 20 min with 10 µL SybrGold in 100 mL 1 × TBE.

5.2.7 Electrospray Ionization Mass Spectrometry (ESI-MS)

The mass spectrometry experiments were run by Adrien Marchand (Dr. Valérie Gabelica's group, University of Bordeaux, France) on an Agilent 6560 DTIMS-Q-TOF instrument from Agilent Technologies (Santa Clara, CA, USA) with a dual-ESI source operated in negative mode. The ESI source voltage was 3.5 kV. The drying gas temperature was adjusted to 200 °C. Soft conditions in the source and the time of flight detector (TOF) were obtained by minimizing the fragmentation of a fragile reference oligonucleotide complex $[(d\underline{G}_4T_4\underline{G}_4)_2 \cdot (NH_4^+)_3]^{5-}$.

The samples were solubilized in nuclease-free grade water from Ambion (Fisher Scientific, Illkirch, France) at approximately 800 µM. The stock concentrations were determined from the absorbance at 260 nm measured on an Uvikon XS spectrophotometer. Ammonium acetate (NH_4OAc, BioUltra for molecular biology, Fluka) was purchased from Sigma-Aldrich (Saint-Quentin Fallavier, France). The injected solutions were obtained by diluting the corresponding volume of the stock and NH_4OAc solutions to reach 5 µM RNA and 100 mM NH_4OAc, respectively. An annealing step was also performed by heating the solutions at 95 °C for 5 min and subsequently allowing them to fast cool on ice.

5.2.8 Dynamic light scattering (DLS)

Measurements were performed in a 12 µL cuvette, dissolving the RNA in water at 0.2-0.24 mM, with addition of either K^+ or Sr^{2+}. Samples were centrifuged before experiments for 30 min at 4 °C and 16'100 × g in order to precipitate any dust that might hamper the measurements. Each sample was recorded at least five times. The arithmetic means of the hydrodynamic radii (r_H) values were subsequently calculated together with the standard deviations. The polydispersity was < 50 % for all measurements.

5.2.9 Nuclear Magnetic Resonance (NMR)

1D-^{1}H NMR

Samples for NMR were generally dissolved in 200-300 µL of 90 % H_2O / 10 % D_2O to allow the detection of imino proton resonances and transferred into 5 mm Shigemi NMR tubes. All samples contained 0.1-0.3 mM RNA and 0-300 KNO_3 and were either prepared at pH 7.4 with addition of 1 mM potassium(I) phosphate buffer or kept in water (measured pH ~ 6.4). Before NMR measurements they were annealed by heating to 95 °C for 5 min and cooling to r.t. for 10 min. 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) was used as an external reference for ^{1}H resonances[354]. 1D-^{1}H-NMR spectra were measured with water suppression at 25 °C unless another temperature is specified. For every measurement condition, two suppression

methods were used to remove the water protons: excitation sculpting (zgesgp) and 11 spin echo (rs_1DH2O), both resulting in similar spectra.

5.2.10 5'-labeling of RNA with an acceptor fluorophore

SCy5 dye (9.15 µL of 8.75 mM stock, 80 nmol, 8 equivalents) was added to 5'-NH_2-NRAS (8.4 µL of 1.2 mM stock, 10 nmol) in 0.1 sodium(I) bicarbonate buffer, pH 8.40. The reaction was allowed to occur overnight at 11 °C and 300 rpm. Subsequently, 8 further equivalents of SCy5 were added to the mixture (9.15 µL of 8.75 mM stock, 80 nmol) for re-labeling under the same conditions. The product was then detected via analytical HPLC and denaturing fluorescent PAGE and purified by HPLC. MALDI-MS confirmed the expected 5'-Cy5-NRAS: MW measured 14808.23 g/mol, 4938.9 m/z $[M-3H]^{3-}$, 7406.5 m/z $[M-2H]^{2-}$, 14806.0 $[M-H]^{-}$; calcd. 4934.7, 7402.6 and 14806.1 m/z, respectively.

5.2.11 smFRET

A microfluidic chamber was self-built from quartz slides obtained from Finkenbeiner (Waltham, MA, USA)[323]. The sample chamber, ~ 8 mm wide and 200 µm deep[228], was created in-house. Two holes were drilled into a quartz slide, to which double-sided tape and a coverslip were added. The edges of the coverslip were fixed with epoxy glue, and the double-sided tape delimited the chamber. Pipetting through the two holes allowed the exchange of the solution during the measurements, in which the observation area was ~ 25 µm x 50 µm[227].

Prior to experiments, the chamber was flushed with 200 µL T50 buffer, followed by the injection of 50 µL biotinylated bovine serum albumin (BSA, 1 mg/mL, 10 % biotinylated), which was allowed to incubate for 10 min. The channel was subsequently flushed with 200 µL T50 buffer and 50 µL of streptavidin solution at 50 µg/mL, followed by 5 min incubation. Excess streptavidin was removed by washing with 100 µL standard buffer (*vide supra*), pH 7.4. This protocol for biotin-streptadivin surface immobilization was adapted from[228,355].

The channel was then flushed with 80 µL RNA:DNA hybrid, previously annealed by heating 5 min at 95 °C and quick-cooling on ice. An enzymatic oxygen scavenger solution (OSS; 165 U/mL glucose oxidase, 1 % D-glucose from Sigma-Aldrich, Buchs, Switzerland, and 2170 U/mL catalase from Thermo Scientific, Rockford, IL, USA) to reduce photobleaching as well as 1 mM Trolox to suppress dye blinking were added. Alternating laser excitation (ALEX) of the green and red lasers was used to check for a constant acceptor signal in order to distinguish between photobleaching events or donor-only labeled molecules and actual conformational changes[332].

Measurements were performed at r.t. upon excitation at 532 nm. Both donor and acceptor emission intensities were monitored over 400 s using a total internal reflection fluorescence

(TIRF) microscope equipped with a charge-coupled device: a 897 CCD camera from Andor (Belfast, UK) allowing 10 frames/s accomplished with an Uniblitz shutter system from Vincent Associates (Rochester, NY, USA) for alternating laser excitation.

5.2.12 smFRET movie processing and data analysis

A composite image was created from each recorded video, in which co-localized spots of SCy3 and SCy5 were identified as single molecules. After corrections for background, bleedthrough (from SCy3 to the SCy5 channel, 7 %), and direct acceptor excitation (green laser excites SCy5 too, by 5 %), the apparent FRET efficiency (E_{app}) was calculated as:

$$E_{\mathrm{app}} = \frac{I_{\mathrm{A}}}{I_{\mathrm{A}} + I_{\mathrm{D}}} \tag{5.5.}$$

where I_A is the fluorescence intensity of the acceptor fluorophore, and I_D, the fluorescence intensity of the donor fluorophore. E_{app} is usually either visualized as a function of time for state determination or histogrammed for thermodynamic analysis.

Time traces of single fluorophore emission were manually selected for bi-labeled molecules using the MATLAB-based software MASH[350].

5.2.13 HPLC

Analytical HPLC

A XTerra RP8 column from Waters (Baden-Dättwil, Switzerland) was used with a flow-rate of 1 mL/min and detection wavelengths 260 and 550 nm. Solvent A was 100 mM triethylammonium acetate (TEAA) pH 7 and solvent B, ACN/H_2O 9:1. The gradient was programmed as follows: 0 min 5 % B, 0-27 min 5-25 % B, 27-33 min 25-95 % B, 33-36 min 95 % B, 36-39 min 95-5% B.

Preparative HPLC

A XTerra RP8 column from Waters (Baden-Dättwil, Switzerland) was used with a flow-rate of 1mL/min and a detection wavelength of 260 nm. Solvent A was 100 mM triethylammonium acetate (TEAA), pH 7, and solvent B, ACN/H_2O 9:1. The gradient was programmed as follows: 0 min 5 % B, 0 27 min 5-15 % B, 27-33 min 15-95 % B, 33-36 min 95 % B, 36-39 min 95-5 % B, 39-40 min 5 % B.

Purities were determined from the HPLC UV/vis traces at 260 and 550 nm. The software for analytical HPLC was Chromeleon 7 from Dionex (Dreieich, Germany).

Chapter 6

Final remarks and outlook

6.1 Final remarks on this work

We have looked at two G-rich RNA sequences postulated to form G4 motifs *in vivo* and having potential as antitumor targets. In *Chapter 2*, we confirm that both NRAS and TERRA fold into very stable parallel G4 motifs, as seen by CD and TDS, and that the structures are already moderately stable in water (T_m ~ 50 ºC), albeit w2ith low folded fractions. Some of the tested monovalent and divalent cations (Na^+, Rb^+, Ca^{2+} and Ba^{2+}; and Mg^{2+} and NH_4^+ only for NRAS) promote further G4 folding by unspecific screening of charge repulsions within the phosphate backbone. Meanwhile, K^+ and Sr^{2+} significantly increase the stability of the G4 folded structures by specifically binding within the channel cavity[285], owing to their compatible size and dehydration energies. As expected, both G4 RNAs are more stable than the analogous DNAs due to the extra 2'-OH groups building further intramolecular contacts. The stabilizing effect of K^+ and Sr^{2+} is comparable for G4 RNAs and DNAs, while a high predilection for K^+ over Na^+ is seen for both RNAs. In contrast, Na^+ stabilization has been reported for DNA to a similar extent than K^+. This difference, with G4 RNAs being stabilized upon K^+ but not Na^+ addition, is presumably related to the fact that G-quadruplexes in RNA are always parallel-stranded, and to the preference of parallel motifs for K^+; while Na^+ shifts G4 DNAs to an antiparallel fold.

Some significant differences exist between the two RNA sequences. For example, the effects of NH_4^+ and Ba^{2+} were only observed on NRAS, which also showed a generally higher sensitivity towards cation addition than TERRA. Both G4s differ in the total number of guanines and these variations might potentially be exploited when designing targeting new strategies for antitumor-relevant G4 RNAs. The systematic study of the cation dependence of the G4 RNAs represents relevant knowledge for the handling of G4 structures *in vivo* and *in vitro* by fundamentally elucidating the best working conditions for such sequences.

In *Chapter 3*, the higher tendency of G4 RNAs towards aggregation is confirmed as opposed to G4 DNAs. The NRAS G4 is observed to multimerize as a G4 dimer in excess of K^+, Sr^{2+}, or NH_4^+ while TERRA remains an intramolecular G4 monomer under the same conditions (*Figure 2*). We explain this difference between the two RNAs by confirming that multimerization is directed by consecutive G bases at the 5' end. These results are to be taken into account to prevent aggregation when designing new short G4 RNA sequences for *in vitro* experiments.

In *Chapter 4*, single-molecule FRET experiments are performed for the NRAS G4 in the presence and absence of K^+ (0-100 mM) and no molecular dynamics are observed. A high FRET state (~ 1) is measured, presumably corresponding to the folded and very stable G4 conformation. We propose an outlook strategy by using a G4-specific G4 resolvase enzyme to unwind the structure and measure its unfolding in real time. This scheme would allow

competition experiments with stabilizing ligands and might give a clear picture of whether the time scale of G4 dynamics in physiological conditions might allow G-quadruplex formation and ligand interaction *in vivo*.

6.2 Outlook on G-quadruplex RNAs

Since the development of the most widely-used G4 prediction algorithm named Quadparser[108], $(\underline{G}_{3+}N_{1-7})_3\underline{G}_{3+}$ has been considered the consensus sequence for G4 formation in nucleic acids. However, in recent years, reports about G4 folding with bulges, *i.e.* non-guanine bases projected out of the G-quartet core and connecting two adjacent guanines of the same column in the G4[356], and long loops (up to 9 nt)[357] indicated that a more relaxed definition should probably be used, such as $(\underline{G}_{3+}N_{1-12})_3\underline{G}_{3+}$[358].

The G4 RNAs known to date include the telomeric transcript TERRA[359] together with several 5' UTR mRNA sequences[148]. All of them are interesting from a targeting point of view due to their antitumor possibilities as translation regulators[165]. 5' UTR sequences that have been confirmed to form G4s comprise NRAS[151], ZIC-1[154], MT3-MMP[360], TRF2[157], CCND3[155], BCL-2[361], ESR1[158] and VEGF[159]. Many others have been postulated from computational predictions and ribosome footprinting[160] studies, such as MYC[160], MYB[160], NOTCH[160], CDK6[160] and ETS1[160], among others. According to the human mRNA sequences from the NCBI's database[362] we postulate here two other potential 5' UTR G4s:

- apoptosis antagonizing transcription factor (AATF)[363]

  5'-GCAGGGAAGGAGCUUCGGGGCCGGGGGUUGGGCCG-3'

- nuclear factor of kappa light polypeptide gene enhancer in B-cells, oncogene (NFKB2) [364]

  5'-CCUGGGUGGCCGGGACAAGAGAAAAGAGGGAGGAGGGCCU-3'

Both of these might potentially act as translation regulators of the corresponding tumor-promoting proteins and are therefore interesting targets for new G4 studies.

So far, no structural data on intramolecular G4 RNAs is reported in the literature. Only a few G4 RNA solution and crystal structures are found in the Protein Data Bank (PDB), including bimolecular TERRA with[174] and without[190,283,285] a G4 acridine ligand. The elucidation of more G4 RNA structures from 5' UTRs or in presence of other ligands is definitely required to further understand the molecular contacts present in these motifs.

The development of selective ligands that are able to distinguish DNA from RNA G4s, or one G4 sequence from another, will be crucial if G4 RNAs are ever to be used as *in vivo* targets for novel therapies. To date, few reports on selective G4 RNA ligands have been

published[172,184]. They are based on specific interactions with the additional 2'-OH group present in RNA, which is forming hydrogen bonds with a negatively-charged functional group from the ligand (e.g. COO^- in carboxypyridostatin)[172]. All G4 RNA sequences characterized so far are folding exclusively as parallel-stranded motifs and thus are monostructural compared to the much more heterogeneous G4 DNAs. Although this limits the structural diversity for selective targeting purposes, the loops in G4 RNAs provide new platforms for ligand interaction as proven by a reported TERRA-acridine complex[174].

Preliminary interaction tests with organometallic complex ligands were undertaken by our collaborators. Dr. Oscar Mendoza (Imperial College London, Prof. Ramon Vilar's lab)* synthesized a series of G4 DNA-stabilizing Ni^{2+} salphen complexes[365]. These complexes are planar and contain two or three arms that will be protonated in physiological conditions. Moreover, their cation center is susceptible to be placed above the G4 ion channel[175]. These type of Ni^{2+} salphen complexes were already reported to bind to G4 DNAs[366,367]. Mendoza prepared the Ni^{2+} compounds OMP_50, OMP_059 with 2 SO_3K, OMP_109 with a 3rd arm, and the SG_43VO V^{4+} complex[365] (*Fig. 6.1.*) and tested them with the 22-nt NRAS sequence 5'-GUGGGAGGGGCGGGUCUGGGUG-3' that we show herein to be monomeric *in vitro* (*Chapter 3*).

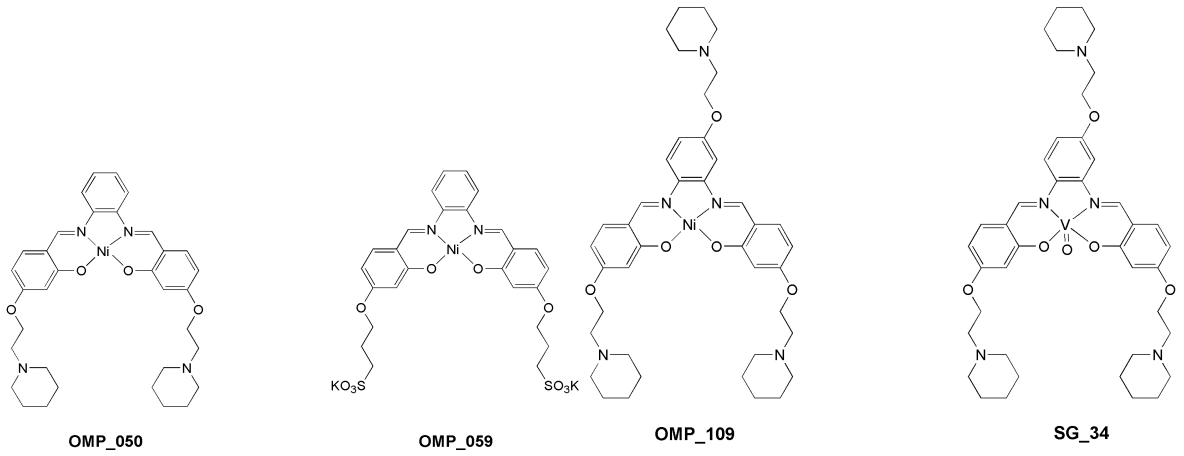


Figure 6.1. Ni^{2+} and V^{4+} salphen complexes. Synthesized by O. Mendoza. The compounds were dissolved in a mixture of DMSO (95% v/v) and 1 mM HCl aqueous solution (5% v/v) to give 5-10 mM stock solutions[365].

A fluorescent intercalator displacement assay (FID) was used by Mendoza to study the interaction of 2.5 µM of each complex with 0.25 µM RNA, in 60 mM potassium(I) cacodylate buffer, pH 7.4. This technique gives a measure of the affinity of each ligand towards the RNA. The compound's concentration at which the fluorescence signal decreases by 50 %,

* Currently working in the lab of Jean-Louis Mergny, University of Bordeaux, France.

corresponding to 50 % displacement of the thyazole orange (TO) dye, is calculated and expressed as a DC_{50} value. The Ni^{2+} salphen complex OMP_059 achieves extremely low displacement of the TO with a DC_{50} value above 4 µM (*Table 6.1.*). This weak binding is expected because of the two sulfite negative charges, which are likely to experience repulsions with the RNA polyanion. On the contrary, the rest of the metal complexes exhibit a considerable displacement of the TO dye, showing DC_{50} values in the range of 0.20 µM (*Table 6.1.*), well below the 0.50 threshold value considered for good G4 binders using this technique[368].

Table 6.1. DC_{50} values (µM) determined using FID assay. Each value is the average of three independent measurements. Results for two G4 DNAs, htelo, and c-myc, are included for comparison and taken from reference[369]. Experiments performed by Mendoza at Imperial College London.

Complex	DC_{50} (µM)		
	NRAS22 RNA	htelo DNA	c-myc DNA
OMP050	0.20	n/a	n/a
OMP059	> 4.0	no binding	no binding
OMP109	0.19	0.26	0.12
SG_34VO	0.20	0.4	0.33

Melting of the G4 RNA was monitored by CD spectroscopy. The negatively charged complex OMP_059 shows very little interaction with the G4 and its presence in the solution leads to a negligible stabilization of the structure (*Table 6.2.*). In contrast, the V^{4+} salphen complex SG_34VO exhibits the highest degree of stabilization, with a ΔT_m practically reaching 30 °C. The triple-substituted complex OMP_109 shows a lower affinity than the double-substituted Ni^{2+} complex OMP_050 (ΔT_m = 19.5 °C and 24 °C, respectively).

Table 6.2. Increase in melting temperature. ΔT_m from CD melting experiments of complex-NRAS22 at 2:1 equivalents. The NRAS22 structure (5 µM) was melted in presence of 2 equivalents of salphen complex in 50 mM Tris-HCl, pH 7.4., and 0.1 mM KCl and the ellipticity was recorded at 260 nm. DNA data from references[365-367,369]. Experiments performed by Mendoza at Imperial College London.

Complex	ΔT_m (°C)		
	NRAS22 RNA	htelo DNA	c-myc DNA
OMP050	24 ± 0.6	33.2±1.0	n/a
OMP059	3 ± 0.2	5.2±1.0	no binding
OMP109	19.5 ± 0.7	26.8±1.5	11.6
SG_34VO	29.8 ± 0.2	6.4±0.9	12

From these results, the V^{4+} salphen complex SG_34VO is the most promising candidate for a G4 RNA stabilizing ligand. Indeed, it presents the highest ΔT_m together with a good affinity and shows some selectivity for RNA over DNA.

During a short-term scientific mission funded by COST Action CM1105, Hannah Pritchard (University of Birmingham, Prof. Michael Hannon's group)† tested in our lab G4 DNA-stabilizing Pt^{2+} and Pd^{2+} compounds for G4 RNA interaction with the TERRA and NRAS sequences[370]. These two square planar complexes, $[Pd(ibiq)_2][BF_4]$ and $[Pt(ibiq)_2][BF_4]$ (ibiq = isoquinoline), present a planar aromatic surface for enhanced π-stacking, a metal centre for electrostatic interactions, and the isoquinoline structure seen in natural medicinal products[371].

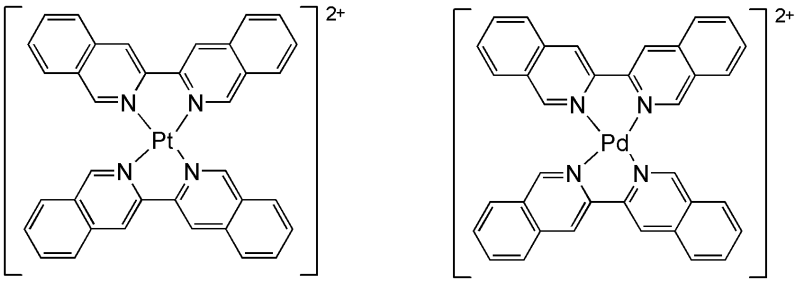


Figure 6.2. Pt^{2+} and Pd^{2+} complexes. $[Pd(ibiq)_2][BF_4]$ and $[Pt(ibiq)_2][BF_4]$, synthesized by Pritchard.

A FID assay was performed in 10 mM sodium(I) cacodylate with 100 mM KCl and 2 equivalents of TO. The obtained DC_{50} values are shown in *Table 6.3*. The Pt^{2+} complex shows a stronger interaction (DC_{50} = 0.2-0.5) than the Pd^{2+} complex with both DNA and RNA. However, no selectivity for DNA or RNA is observed. The isoquinoline rings might be functionalized in the future by including side arms that might then target either DNA (with protonated groups) or RNA (with H bond forming groups).

Table 6.3. DC_{50} values (μM) determined using the FID assay. Two DNA sequences, htelo and c-myc, are included for comparison[370]: htelo22 5'-AGGGTTAGGGTTAGGGTTAGGG-3' and cmyc22 5'-TGAGGGTGGGTAGGGTGGGTAA–3'. Experiments performed by Pritchard at the University of Birmingham and in our lab.

Complex	DC_{50} (μM)			
	htelo22 DNA	c-myc DNA	TERRA RNA	NRAS18 RNA
Pd ibiq	0.5	0.3	> 2.5	1.9
Pt ibiq	0.5	0.2	0.3	0.4

As described above, the design of a suitable metal complex as a ligand for G4 RNA is not trivial. Research into specific ligands will have to continue, if ever these structures are to be exploited as *in vivo* targets. In conclusion, G4 RNA sequences are very stable motifs in physiological conditions, and their high conservation and affinity for K^+ suggests that they are

† Currently working as Assistant Quality Manager at Cambridge Commodities Ltd, UK.

playing a role in the cytoplasm of living cells (*Fig. 6.3.*), in which the K^+ concentration is ca. 140 mM.

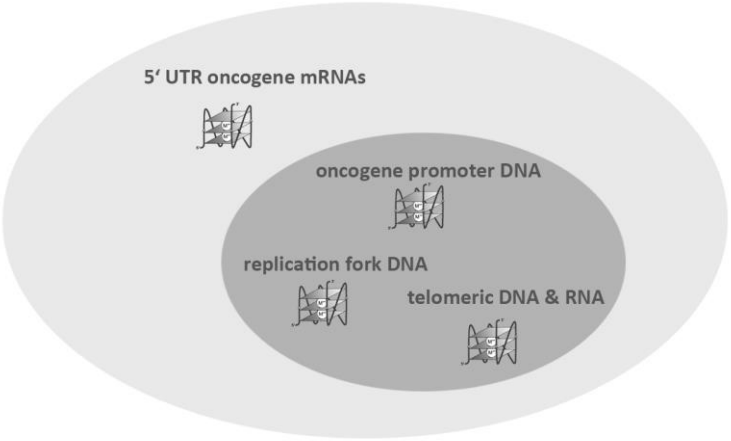


Figure 6.3. G4 motifs in the cell. Location of G-quadruplex structures postulated to form in living cells and to have regulatory roles *in vivo* [132].

Current challenges in the field lie in obtaining more structural data, finding selective ligands, and understanding the interaction with G4-binding proteins that potentially have a G4-regulatory action *in vivo*. Indeed, in 2014, Gray *et. al.* proposed that the elucidation of more endogenous proteins binding G4 DNA would allow to establish a functional map of the G4 genome[372]. In a similar fashion, obtaining a more accurate picture of the G4 transcriptome would give valuable information related to the ncRNA-mediated regulation of biological processes.

In less than 60 years since F. Crick postulated the existence and function of mRNA, a whole new “RNA world” has emerged. Nowadays, it is known that RNA comes in many forms (*Fig. 6.4.*) and carries out many different essential functions in living organisms. The current challenge is to expand and exploit this knowledge for the development of RNA-based therapeutic strategies that might either allow the treatment of so-far-untreatable conditions, or overcome the side-effects of currently used drugs. The fact that G4 RNA structures appear to be involved in a wider range of biological processes, compared to their DNA counterparts (*Fig. 6.5.*), might make it possible to validate some of them as therapeutic targets.

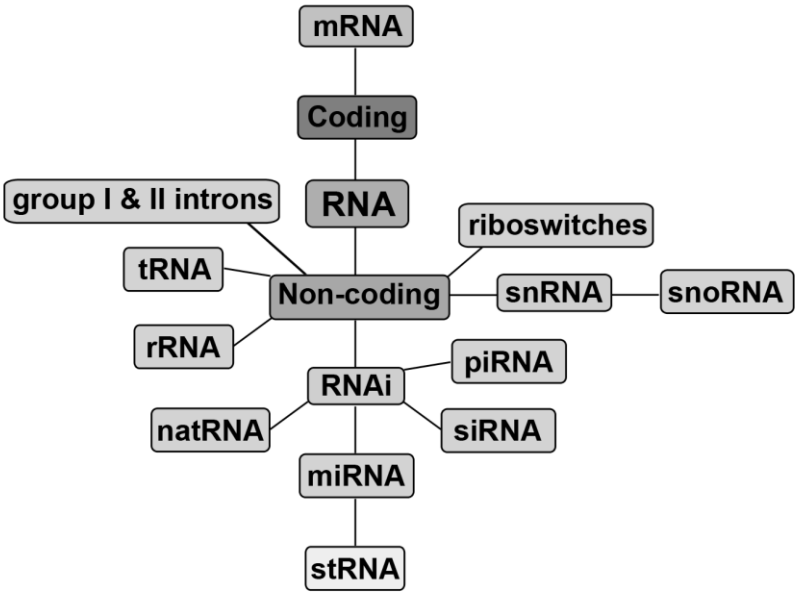


Figure 6.4. Types of biological RNAs and classification into coding and non-coding. Messenger RNA (mRNA) is known as coding RNA due to its function as intermediate genetic code carrier for the synthesis of proteins. All other classes of RNA belong to the non-coding family; *i.e.*, their sequence is not expressed in the form of a protein. Rather, they play diverse key roles in living cells, including the catalysis of protein synthesis (rRNA), amino acid transfer (tRNA), transcriptional (riboswitches) and post-transcriptional regulation (riboswitches, RNAi), rRNA modification (snoRNA), germline development regulation (piRNA), and self-splicing (group I and II introns). G-quadruplexes are thought to fold in non-coding regions of the RNA, including 5' and 3' UTRs of mRNA and rRNA, among others.

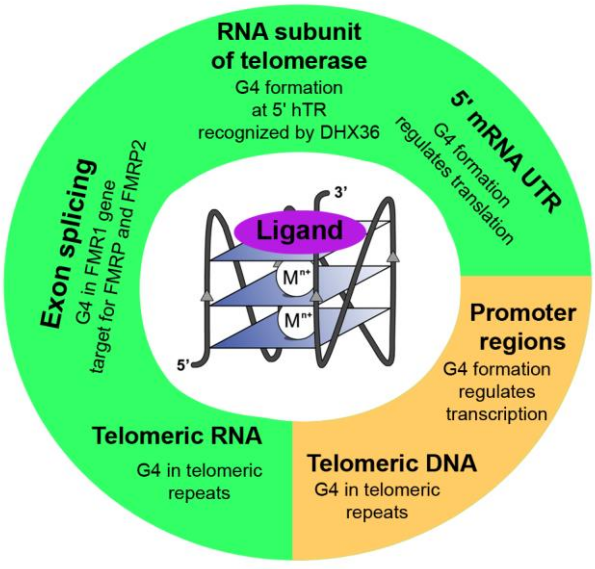


Figure 6.5. DNA and RNA G4 structures within the human genome and transcriptome. These motifs all have potential to be targeted with small molecule ligands for the development of new therapeutic strategies. Adapted from reference[165].

7. Summary

Guanine-rich nucleic acid sequences have the tendency to fold into non-canonical helical structures, known as G-quadruplexes (G4s). These consist of a stack of two or more G-quartets, each of which is a square-planar arrangement of hydrogen-bonded guanine bases. G4 RNAs are found in regulatory regions of the transcriptome, especially in 5' untranslated regions (UTRs), and can be detected in the cytoplasm of immobilized human cells. They have a very high thermodynamic stability in physiological conditions and a conserved parallel-stranded fold.

We are interested in G4 RNAs as relevant regulatory elements present *in vivo* and serving as potential targets for novel therapies, e.g. antitumor therapies. Indeed, the extra 2'-OH only present in RNA has been shown to allow the design and development of small molecules able to selectively target G4 RNAs over G4 DNAs. We work with the 18-nt sequence 5'-GGGAGGGGCGGGUCUGGG-3' from the 5' UTR of the NRAS oncogene mRNA (NRAS stands for neuroblastoma rat sarcoma viral oncogene homolog), which codes for a cellular-proliferation-related protein. When folded into a G4, NRAS has previously been shown to reduce protein synthesis *in vitro* by translation inhibition and can be targeted with a small ligand, thus making it interesting for novel anticancer efforts. We also use the well-documented telomeric repeat-containing RNA (TERRA), the transcript of the telomeric DNA repeat. It was shown that the 24-nt long sequence 5'-$(UUAGGG)_4$-3' folds into a stable G4 in the presence of K^+ and inhibits the activity of telomerase. This enzyme is overexpressed in ca. 80 % of human cancers and is responsible for tumor cell immortality. TERRA, as well as its G4 DNA counterpart, is therefore being investigated as a potential antitumor target, as the TERRA G4 would have a telomerase-inhibition-mediated anticancer effect if stabilized *in vivo*.

Like all DNAs and RNAs, G4s have a high negative charge due to their phosphate backbone and are therefore sensitive to the surrounding cationic conditions. Moreover, the oxygen lone pairs from the carbonyl groups of the G-quartet-forming guanines give rise to a locally high negative charge density at the center of the planes thus forming a specific cation binding site able to coordinate metal ions, which are needed for the folding and stability of the G4 structures (*Fig. 7.1.*). As a consequence of this defined channel binding site only cations with an adequate size, charge, and dehydration energy are predicted to specifically interact with G4s.

Several reports have been published on the interaction between G4 DNA and metal ions while less information is available on G4 RNAs. They are known to be more stable than their DNA counterparts but no systematic studies with metal ions have been reported. Therefore, we aim at understanding metal ion interactions in G4 RNAs and the cation effect on their folding and

stability, while comparing the results with the known data on G4 DNAs. We also want to observe metal ion-dependent dynamics at the single-molecule level in order to elucidate the steps of G4 RNA folding.

Chapter 2 aims at throwing light on the effect of monovalent (Li^+, Na^+, K^+, Rb^+, Cs^+, and NH_4^+) and divalent (Mg^{2+}, Ca^{2+}, Sr^{2+}, and Ba^{2+}) metal ions on the folding and stability of G4 RNAs. We use circular dichroism and thermal difference UV spectra to verify G-quadruplex formation of NRAS and TERRA, which are shown to fold into parallel G4s. This is followed by further circular dichroism as well as UV thermal melting experiments to assess the degree of folding and relative stability, respectively, in each of the ionic conditions tested. The circular dichroism results obtained in water reveal that the G4s are pre-folded without cation addition with some CD signal arising from guanine pre-stacking. Subsequent addition of mono- or divalent metal ions increases the folded G4 form, as seen by an increase in intensity of the CD signals. Specifically, Na^+, K^+, Sr^{2+}, and Rb^+ promote the folding in both NRAS and TERRA sequences, while Ba^{2+} and NH_4^+ favor NRAS folding only. At a higher concentration, Ca^{2+} also increases G4 folding for both RNAs, as does Mg^{2+} for NRAS.

At the same time, only K^+ and Sr^{2+} result in a significant increase of the melting temperatures of both NRAS and TERRA, which are already moderately stable in water-only conditions (T_m's ~ 50 °C). This higher stabilization of the G4 motifs in presence of K^+ and Sr^{2+} is already known for G4 DNA and is explained by the optimal size and dehydration energies of these two cations, allowing them to fit in the G4 cation binding channel between the G-quartets. In contrast to DNA, there was no further stabilization observed by addition of Na^+ even at high concentrations up to 100 mM. This is presumably due to the preference of parallel G4s for K^+ binding and the fact that G4 RNAs are always parallel-stranded. In contrast, G4 DNAs are more heterogeneous in conformation, folding also antiparallelly or in hybrid forms depending on the surrounding conditions.

We have shown the cation dependence of two parallel-only G4 RNAs and compared it to the reported cation interaction of G4 DNAs. The main difference is the Na^+ insensitivity of the RNA motifs' stability, together with the fact that the G4 RNAs are already quite stable in water-only conditions, in contrast to their DNA counterparts. Moreover, we have observed some differences in cation interaction for both sequences: the NRAS RNA is more sensitive to cation addition as manifested by the influence of Ba^{2+}, NH_4^+ and Mg^{2+} on its folding. The fact that these cations have a significant effect on NRAS and no measurable effect on TERRA hints that differences in cation response might potentially be exploited for the selective targeting of a specific G4 RNA. All in all, the systematic study of G4 RNAs' cation dependence underlines the importance of elucidating the best working conditions for such sequences, enabling researchers to control G4 folding and tune G4 stability.

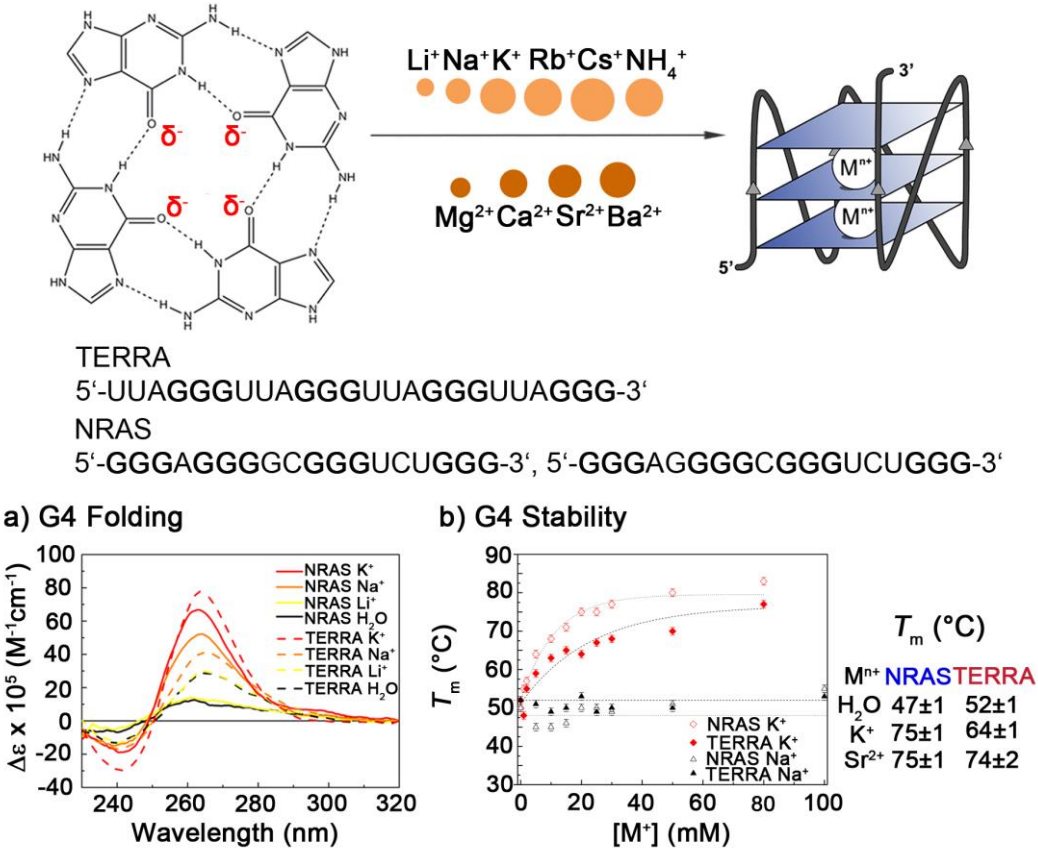


Figure 7.1. Interaction of G-quadruplex RNAs (TERRA and NRAS) with monovalent and divalent metal ions (Li^+, Na^+, K^+, Rb^+, Cs^+, and NH_4^+; and Mg^{2+}, Ca^{2+}, Sr^{2+}, and Ba^{2+}). NRAS can form two different G4 motifs, depending on which Gs are contained in the G-quartet. a) CD spectra in water-only shows that Na^+ and K^+, among others (*not shown here*), promote G4 folding; b) K^+ increases the G4 stability of both NRAS and TERRA while Na^+ has no effect. Indeed, K^+ and Sr^{2+} are the only two cations that yield a significant increase of the RNA melting temperatures (T_m), compared to water-only controls.

The spectroscopic studies described in Chapter 2 give some indications of possible G4 multimerization or aggregation, especially for NRAS. For example, the melting temperatures are not independent of the NRAS concentration above 10-15 µM, indicating that G4 multimers are probably present in solution. G4 multimerization is already reported in the literature for short DNA and RNA sequences, especially when these contain short loops or consecutive guanines at the ends.

Chapter 3 therefore dwells into the multimerization tendency of the NRAS G4 in high excess of stabilizing cations. This RNA sequence forms a parallel G4 unit that easily multimerizes in presence of K^+, Sr^{2+} or NH_4^+. Native polyacrylamide gel electrophoresis (PAGE) in the mentioned cation conditions presents a slow-running band that might correspond to an NRAS dimer. This G4 dimer is subsequently confirmed by mass spectrometry in soft ionization conditions (ESI-MS). In contrast, the TERRA G4 appears as a monom2er both in PAGE and MS experiments. Determination of the hydrodynamic radii (r_H) in K^+ or Sr^{2+} solution is carried

out via dynamic light scattering (DLS) and confirms the appearance of higher multimerization states at higher mono- or divalent cation concentrations for both RNAs. Moreover, r_H values for NRAS correspond again with a higher tendency towards multimerization for this sequence.

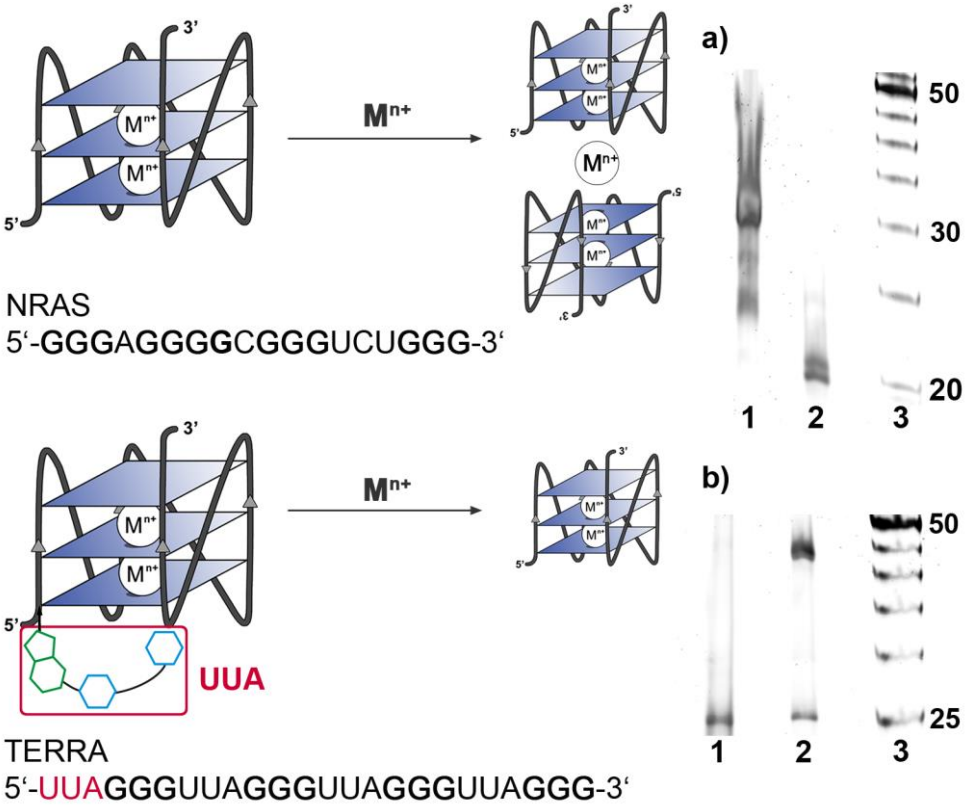


Figure 7.2. Multimerization behavior of G-quadruplex RNAs (NRAS and TERRA) in excess of stabilizing cations (K^+, NH_4^+, Sr^{2+}). The effect of the flanking sequences is observed, resulting in 24-nt TERRA being monomeric, while 18-nt NRAS dimerizes. a) Native PAGE in 20 mM K^+ showing NRAS18, dimer (1), and NRAS21, monomer (2); b) Native PAGE in 20 mM K^+ showing TERRA24, monomer (1), and TERRA21, mixture of monomer and dimer (2).

In order to understand the difference in multimerization tendency between NRAS and TERRA, flanking nucleotides from the natural mRNA sequence were added to NRAS and the multimerization tendency of the modified sequences was explored by PAGE and ESI-MS. Addition of one flanking nucleotide at each end (NRAS20) results in dimerization. In contrast, two added bases at each the 5' and the 3' ends (NRAS22) hinder the multimerization, yielding an NRAS monomer. Moreover, an NRAS to which a flanking UUA sequence at the 5' end has been attached (NRAS21) does not aggregate any more. At the same time, deleting the first three nucleotides (UUA) from the TERRA sequence yields a shorter G4 (TERRA21) that dimerizes. These PAGE findings suggest that the reason for the original NRAS multimerization is the presence of three consecutive Gs at the 5' end, and that the G4 dimer presumably forms via 5'-5' stacking. MS experiments confirm that the NRAS dimer contains six quartet planes and five coordinated metal ions (*Fig. 7.2.*).

We corroborate that G4 RNAs have a higher tendency than G4 DNAs towards multimerization. This is possibly due to the parallel RNA form allowing better π-stacking of G4 monomers. The NRAS G4 multimerizes as a G4 dimer in K^+, Sr^{2+}, or NH_4^+ while TERRA remains an intramolecular G4 monomer under the same conditions (*Fig. 7.2.*). We have been able to explain this difference between the two RNAs by measuring PAGE and MS with modified G4 sequences and proving that multimerization is directed by the 5' end containing consecutive, G4-forming, Gs. These results are to be taken into account to prevent aggregation when designing new short G4 RNA sequences for *in vitro* experiments.

More than only one folded and one unfolded G4 form can be present in solution if the G4 RNA is aggregating, yielding an heterogeneous average in both the spectroscopic and gel experiments discussed in Chapters 2 and 3. Together with potential folding intermediates upon cation titration, this will give rise to heterogeneities, which are not observable in the bulk measurements described so far. Aiming at observing heterogeneities in the G4 RNAs folding paths or in their multimerization steps, single-molecule (sm) experiments were undertaken with the use of Förster resonance energy transfer (FRET).

Chapter 4 thus explores the application of single-molecule FRET to the study of G4 RNA. A working construct for the purpose of FRET studies is designed with a commercial NRAS sequence, which is labeled in house with a SCy5 fluorophore at the 5' end, via NHS-ester coupling in buffered aqueous conditions. The successful labeling is shown by mass spectrometry of the SCy5-NRAS RNA. Using a commercial DNA oligonucleotide, internally labeled with SCy3 and containing biotin at the 5' end, the RNA:DNA hybridization conditions are optimized. Immobilization of the resulting construct onto quartz slides is subsequently achieved via biotin-streptavidin linkage and thus a working system for single-molecule observation of the NRAS G4 is setup.

Single-molecule FRET movies are recorded in the presence and absence of K^+ (0 – 100 mM), using alternating laser excitation (ALEX) to select molecules containing both fluorophores. Some anticorrelation of donor with acceptor fluorescence is observed upon SCy5 bleaching (*Fig. 7.3.*) indicating that FRET occurs between the donor and the acceptor dyes. However, no transitions between FRET states are present and only one high FRET state (~ 1) is seen. Presumably, the NRAS G4 stays in a stable folded conformation and thus no folding intermediates are seen. This is not surprising considering the high NRAS RNA stability reported in Chapter 2. At the same time no NRAS multimerization is observed at the single-molecule level. On the one hand, low concentrations in the pM range and the presence of a bulky SCy5 fluorophore at the 5' end of the RNA might hinder aggregation during sample preparation, and on the other hand the single molecule immobilization to a surface prevent RNA-RNA interactions from appearing after the excess NRAS is washed away.

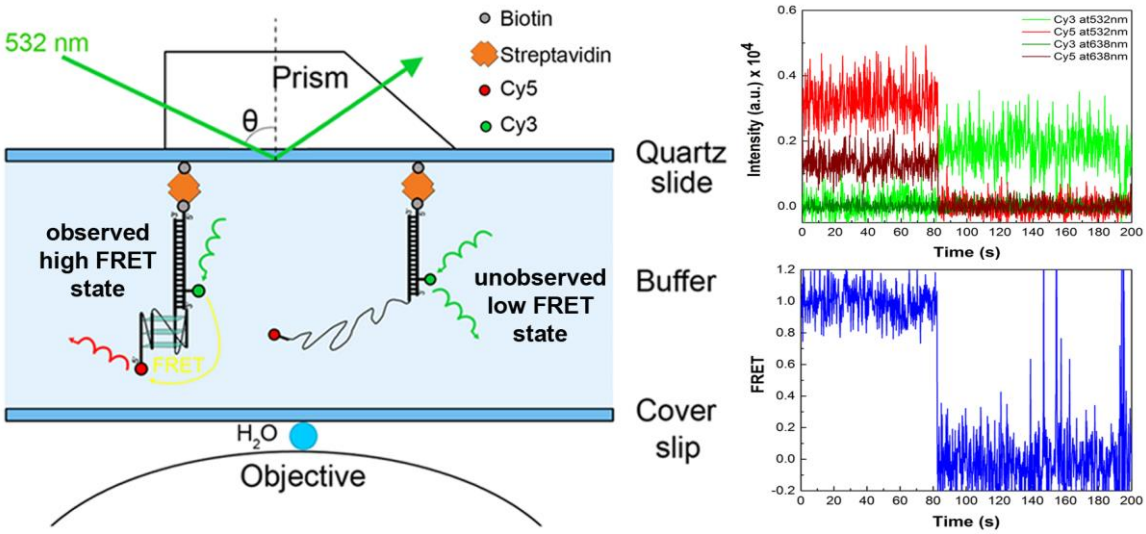


Figure 7.3. Single-molecule FRET measurements of the NRAS G-quadruplex. Only the high FRET state, corresponding to a folded and stable G4, is observed. An example of a typical time trace is shown together with the respective apparent FRET value. In this molecule, measured without K^+ addition, SCy5 bleaching occurs at 80 s and anticorrelation is observed between donor and acceptor fluorescence. This confirms that FRET is occurring between both dyes.

The chosen strategy for single-molecule observation of the NRAS G4 yields one stable G4 folded conformation with a high FRET value. No dynamics or intermediate FRET states are observed even without K^+ cation present in the imaging buffer. We therefore propose an outlook strategy which would use the G4-specific G4 resolvase enzyme to unwind the structure, thus allowing to measure its unfolding in real time. This scheme would allow competition experiments with stabilizing ligands and might give a clear picture of whether the time scale of G4 dynamics in physiological conditions would allow G-quadruplex formation and ligand interaction *in vivo*.

Overall, this thesis looks at the *in vitro* behavior of two short G-rich RNA sequences known to fold into G-quadruplex motifs. Their metal ion dependency is elucidated together with their tendency towards multimerization and both aspects are compared to the existent knowledge on G4 DNAs. In this way, a good basis for the best buffer conditions and the choice of monomeric RNA sequences for *in vitro* experiments is established. The observations of the NRAS G4 at the single-molecule level further emphasize that these G4 RNAs are very stable, so much so that no dynamics are measured, neither in presence nor in absence of K^+, in the time frame of the experiments. The herein reported results represent another step towards the understanding of these RNA motifs, which are relevant for *in vivo* regulation of biological processes and have potential as future antitumor targets.

8. Zusammenfassung

Guanin-reiche Nukleinsäuresequenzen haben die Tendenz sich in nicht-kanonische helikale Strukturen zu falten, sogenannte G-Quadruplexe (G4). Diese bestehen aus einem Stapel von zwei oder mehr G-Quartetten, jedes davon in einer quadratisch-planaren Anordnung von Wasserstoff-gebundenen Guaninbasen. G4 RNS findet man in regulierenden Regionen des Transkriptoms, insbesondere dem 5‘ untranslatierte Regionen (UTRs), und sie können im Zytoplasma von immobilisierten menschlichen Zellen detektiert werden. Sie haben eine sehr hohe thermodynamische Stabilität unter physiologischen Bedingungen und eine konservierte parallel-strängige Faltung.

Wir sind an G4-RNS sowohl als regulatorische Elemente, die *in vivo* vorkommen, als auch als potentielle Ziele für neue Therapien, z.B. Antitumortherapien, interessiert. In der Tat hat die zusätzliche 2'-OH-Gruppe, die nur in RNS präsent ist, gezeigt, dass das Design und die Entwicklung von kleinen Molekülen möglich sind, die in der Lage sind, selektiv auf G4-RNS und nicht auf G4-DNS abzuzielen. Wir arbeiten mit der 18-Nulkeotidsequenz 5'-GGGAGGGGCGGGUCUGGG-3' aus der 5‘ UTR der NRAS Onkogen-mRNS (NRAS steht für Neuroblastoma Rat Sarcoma viral oncogene homolog), welche ein auf Zellvermehrung bezogenes Protein kodiert. Für NRAS wurde bereits gezeigt, dass es gefaltet zu einen G4 die Proteinsynthese *in vitro* durch die Translationsinhibierung reduziert, und dass es mit einem kleinen Liganden angesteuert werden kann. Dies macht es interessant für neue Antikrebs Strategie. Wir verwenden auch die gut-dokumentierte, telomerische Wiederholungen enthaltende (TERRA) Sequenz, das Transkript der telomerischen DNS Wiederholung. Es wurde gezeigt, dass die 24-Nukleotide lange Sequenz 5'-(UUAGGG)$_4$-3' in Anwesenheit von K^+ in eine stabile G4 faltet und, dass sie die Aktivität der Telomerase inhibiert. Dieses Enzym wird bei ca. 80% der menschlichen Krebse überexprimiert und ist verantwortlich für Tumorzellenunsterblichkeit. TERRA-G4 und dessen Gegenpart werden deshalb als potentieller Antitumorziele untersucht, da TERRA-G4 einen telomeraseinhibierenden Antikrebseffekt hätte falls es *in vivo* stabilisiert werden kann.

Wie alle DNS und RNS haben G4 eine hohe negative Ladung aufgrund ihres Phosphatzuckerrückgrats und sind deshalb sensitiv gegenüber den vorliegenden/umgebenden Kationenbedingungen. Darüber hinaus führen die Sauerstoffelektronenpaare der Carbonylgruppen der G-Quartett-bildenden Guanine zu einer lokal erhöhten negativen Ladungsdichte am Zentrum der Fläche und bilden dadurch eine spezifische Kationenbindungsstelle, die in der Lage ist Metallionen zu koordinieren, welche für die Faltung und Stabilität der G4-Strukturen benötigt werden (*Abb. 8.1.*). Als Konsequenz dieser definierten kanalartigen Bindungsstelle sollten nur Kationen mit einer angemessenen Grösse, Ladung und Dehydratationsenergie spezifisch wechselwirken können mit G4en.

Es wurden einige Berichte zu der Wechselwirkung zwischen G4 DNS und Metallionen publiziert, währenddessen es zu G4 RNS weniger Informationen gibt. Letztere sind bekanntermassen stabiler als ihre DNS Gegenspieler, aber es sind keine systematischen Studien mit Metallionen verzeichnet. Deshalb zielen wir auf das Verständnis von Metallion Wechselwirkungen mit G4 RNS und den Kationeneffekt auf die Faltung und Stabilität ab im Vergleich der Ergebnisse mit den bekannten Daten zu G4 DNS. Wir wollen zudem Metallion-abhängige Dynamiken auf dem Einzelmoleküllevel beobachten um die Schritte der G4 RNS Faltung aufzuklären.

Kapitel 2 zielt auf die Auflösung des Effekts monovalenter (Li^+, Na^+, K^+, Rb^+, Cs^+ und NH_4^+) und bivalenter (Mg^{2+}, Ca^{2+}, Sr^{2+} und Ba^{2+}) Metallionen auf die Faltung und Stabilität von G4 RNS ab. Wir verwenden zirkularen Dichroismus und thermische Differenz UV Spektren um die G-Quadruplexbildung von NRAS und TERRA zu verifizieren, für welche gezeigt wurde, dass sie in parallele G4s falten. Dies wird mittels weiter Dichroismus sowie UV thermischen Schmelzexperimenten verfolgt um jeweils den Grad der Faltung und der relativen Stabilität in jeder der getesteten ionischen Bedingungen zu ergründen. Die zirkularen Dichroismus-Ergebnisse in Wasser zeigen dass die G4 ohne Zugabe von Kationen vorgefaltet sind, wobei manche CD Signale von der Guanin Vorfaltung kommen. Anschliessende Zugabe von mono- oder bivalenten Metallionen erhöhen die gefaltete G4-Form danach, erkennbar an einer Erhöhung der Intensität der CD Signale. Vor allem Na^+, K^+, Sr^{2+} und Rb^+ unterstützen die Faltung sowohl von NRAS als auch TERRA Sequenzen, während Ba^{2+} und NH_4^+ nur NRAS Faltung begünstigen. Bei höheren Konzentrationen erhöht auch Ca^{2+} die G4 Faltung für beide RNS, ebenso wie auch Mg^{2+} für NRAS.

Zugleich führen nur K^+ und Sr^{2+} zu einem signifikanten Anstieg der Schmelztemperaturen von NRAS wie auch TERRA, welche bereits in nur-Wasser Bedingungen einigermassen stabil sind (T_m's ~ 50 °C). Diese höhere Stabilisierung des G4-Motivs in K^+ und Sr^{2+} ist bereits für G4 DNS bekannt und erklärt sich durch die optimale Größe und Dehydratationsenergie dieser beiden Kationen, die es ihnen erlauben in den G4 Kationenbindungskanal zwischen den G-Quartetten zu passen. Im Gegensatz zu DNS wurde keine weitere Stabilisierung durch die Zugabe von Na^+ beobachtet, selbst bei hohen Konzentrationen von bis zu 100 mM. Dies ist wahrscheinlich auf die Präferenz paralleler G4 für die K^+-Bindung und der Tatsache dass G4 RNS immer parallelsträngig sind zurückzuführen. Sttattdessen sind mit antiparallelen oder Hybridform-Faltungen G4 DNS heterogener in ihrer Konformationen, welche von den umgebenden Kationen abhängen.

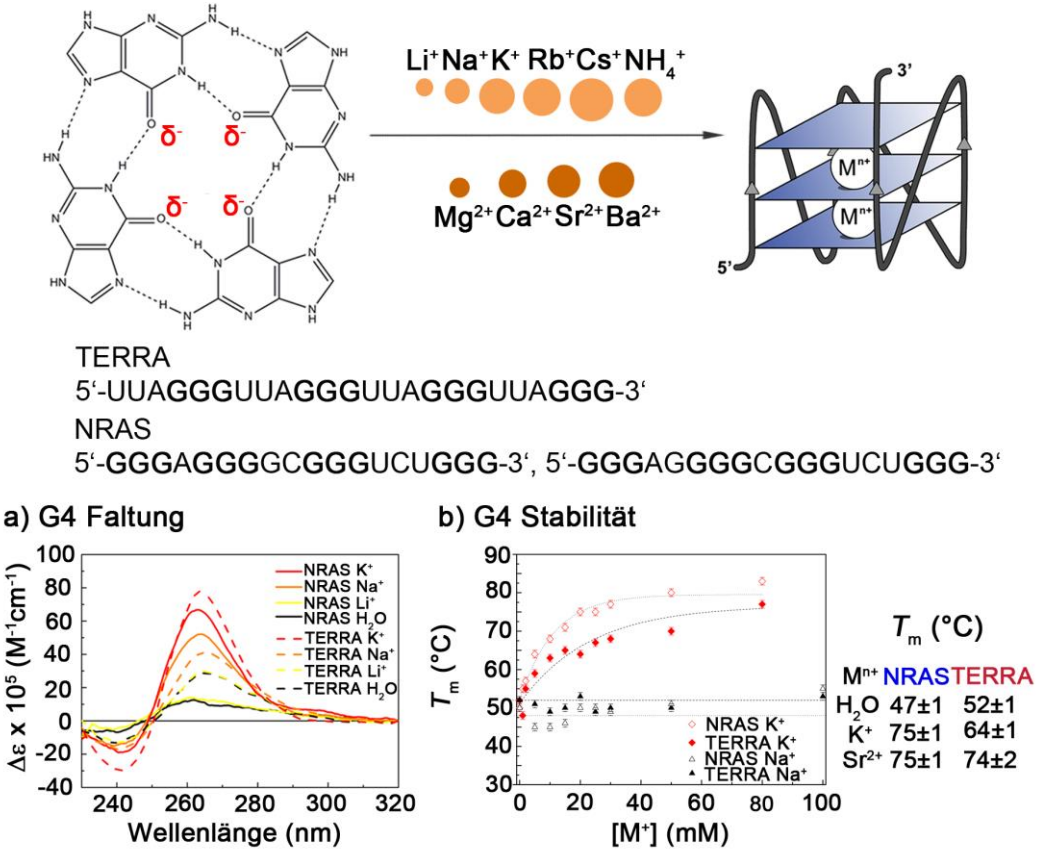


Abbildung 8.1. Wechselwirkung von G-Quadruplex RNS (TERRA und NRAS) mit monovalenten und divalenten Metallionen (Li^+, Na^+, K^+, Rb^+, Cs^+ und NH_4^+; Mg^{2+}, Ca^{2+}, Sr^{2+} und Ba^{2+}). NRAS kann zwei verschiedene G4 Motive bilden, abhängig davon welche G im G-Quartett enthalten sind. a) CD in reinem Wasser zeigt, dass Na^+ und K^+ (neben anderen, *hier nicht gezeigt*) die G4 Faltung unterstützen; b) K^+ erhöht die G4 Stabilität von NRAS wie auch TERRA, während Na^+ keinen Effekt hat. In der Tat sind K^+ und Sr^{2+} die einzigen zwei Kationen, die zu einem signifikanten Anstieg der RNS Schmelztemperaturen (T_m) führen.

Wir haben die Kationen Abhängigkeit von zwei ausschliesslich parallelen G4 RNS untersucht und mit der berichteten Kationenwechselwirkung von G4 DNS verglichen. Der Hauptunterschied ist die Na^+-Insensitivität der RNS Motivstabilität zusammen mit der Tatsache, dass die G4 RNS bereits in reinen Wasser ziemlich stabil sind, jedenfalls mehr als ihr Gegenspieler DNS. Darüber hinaus haben wir gewisse Unterschiede in der Kationenwechselwirkung beider Sequenzen beobachtet: die NRAS Sequenz ist sensitiver gegenüber Kationenzugabe, unterstrichen durch den Einfluss von Ba^{2+}, NH_4^+ und Mg^{2+} auf ihr Faltung. Die Tatsache dass diese Kationen einen signifikanten Effekt auf NRAS haben aber keinen messbaren Effekt auf TERRA deutet an, dass Unterschiede in der Koordination von Kationen möglicherweise für das selektive Targeting von spezifischen G4 RNS ausgenutzt werden könnten. Zusammengefasst unterstreicht diese systematische Studie der G4-RNS-Kationenabhängigkeit die Wichtigkeit die besten Arbeitsbedingungen für solche Sequenzen zu finden, was es Wissenschaftlern ermöglicht *in vitro* die G4 Faltung zu kontrollieren und die G4 Stabilität zu tunen.

Die in Kapitel 2 beschriebenen spektroskopischen Studien geben insbesondere für NRAS etliche Hinweise auf mögliche G4 Multimerisierung und Aggregierung. Zum Beispiel sind die Schmelztemperaturen nicht unabhängig von der NRAS Konzentration über 10-15 µM, was darauf hinweist, dass auch intermolekulare Spezies in Lösung anwesend sind. G4 Multimerisierung ist bereits in der Literatur für kurze DNS und RNS Sequenzen beschrieben, besonders wenn diese kurze Schlaufen oder konsekutive Guanine am Ende enthalten.

Kapitel 3 geht deshalb näher auf die Tendenz zur Multimerisierung von NRAS G4 bei hohem Überschuss von stabilisierenden Kationen ein. Diese RNS Sequenz bildet eine parallele G4 Einheit, die leicht in Gegenwart won K^+, Sr^{2+} oder NH_4^+ multimerisiert. Native Polyacrylamid Gelelektrophorese (PAGE) unter den erwähnten Kationenbedingungen zeigt eine langsam laufende Bande die zu einem NRAS Dimer gehören könnte. Dieses G4 Dimer wird danach durch Massenspektrometrie unter milden Bedingungen (ESI-MS) nachgewiesen. Im Gegensatz dazu erscheint TERRA G4 als Monomer sowohl in PAGE als auch im MS. Die Bestimmung der hydrodynamischen Radien (r_H) in Gegenwart von K^+ oder Sr^{2+} wird mittels dynamischer Lichtstreuung (DLS) durchgeführt, was die Anwesenheit von höheren Multimerisierungszuständen bei höheren mono- oder bivalenten Kationenkonzentrationen für beide RNS bestätigt. Darüber hinaus stimmen die r_H Werte für NRAS wieder mit der höheren Tendenz gegenüber Multimerisierung für diese Sequenz überein.

Um den Unterschied in der Multimerisierungstendenz zwischen NRAS und TERRA zu verstehen, wurden flankierende Nukleotide der natürlichen mRNS Sequenz zu NRAS hinzugefügt und die Multimerisierungstendenz der modifizierte Sequenzen wiederum mittels PAGE und ESI-MS erforscht. Die Zugabe von einem flankierenden Nukleotid an jedem Ende (NRAS20) resultiert immer noch in Dimerisierung. Im Gegensatz dazu verhindern zwei zugegebene Basen jeweils am 5‘- und am 3‘-Ende (NRAS22) die Multimerisierung, was zu einem NRAS Monomer führt. Darüber hinaus aggregiert NRAS zu dem flankierend UUA am 5‘-Ende angefügt wurde (NRAS21) nicht mehr. Gleichzeitig ergibt die Entfernung der ersten drei Nukleotide (UUA) aus der TERRA Sequenz eine kürzere G4 (TERRA21), die dimerisiert. Diese Erkenntnisse legen nahe, dass der Grund für das originale NRAS Dimer welches sich bevorzugt über 5‘-5‘ Stapelung bildet die Abwesenheit eines 5‘-Überhangs ist. MS bestätigt, dass das NRAS Dimer sechs Quartettebenen und 5 koordinierte Metallionen enthält (*Abb. 8.2.*).

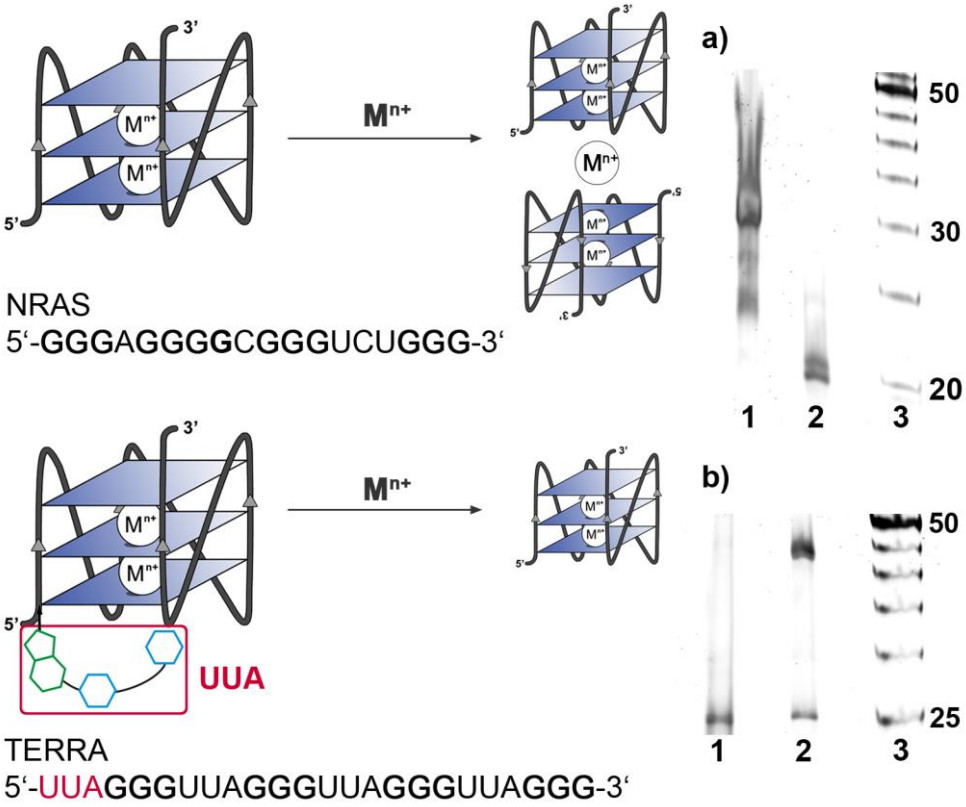


Abbildung 8.2. Multimerisierungsverhalten von G-Quadruplex RNS (NRAS und TERRA) im Überschuss von stabilisierenden Kationen (K^+, NH_4^+, Sr^{2+}). Der Effekt von flankierenden Sequenzen wird untersucht, was in eine 24-nt monomere TERRA Sequenz resultiert, während 18-nt NRAS dimerisiert. a) Native PAGE in 20 mM K^+ zeigt NRAS18 Dimer (1) und NRAS21 Monomer (2); b) Native PAGE in 20 mM K^+ zeigt TERRA24 Monomer (1) und ein TERRA21 Gemisch von Monomer und Dimer (2).

Wir bestätigen, dass G4 RNS eine höhere Tendenz als G4 DNS zur Multimerisierung haben. Dies ist möglicherweise aufgrund der parallelen RNS Form, die ein besseres π-Stacking der G4 Monomere erlaubt. Die NRAS G4 multimerisiert als G4 Dimer in K^+, Sr^{2+} oder NH_4^+; während TERRA ein intramolekulares G4 Monomer unter den gleichen Bedingungen bleibt (*Abb. 8.2.*). Wir waren in der Lage diesen Unterschied zwischen den zwei RNS zu erklären durch das Messen von PAGE und MS mit modifizierten G4 Sequenzen und das Nachweisen, dass die Multimerisierung durch das 5'-Ende gelenkt wird, das konsekutive G4-bildende, G enthält. Diese Ergebnisse müssen berücksichtigt werden um Aggregation zu verhindern wenn neue kurze G4 RNS Sequenzen für *in vitro* Experimente entworfen werden.

Mehr als nur eine gefaltete und eine ungefaltete G4-Form kann in Lösung anwesend sein falls die G4 RNS aggregiert, was ein heterogenen Durchschnitt ergibt in spektroskopischen Experimenten, die in Kapitel 2 und 3 durchgeführt wurden. Dies führt zusammen mit potentiellen Faltungsintermediaten während Kationentitration zur Erhöhung von Heterogenitäten, welche nicht in den bis jetzt beschriebenen Bulkmessungen beobachtbar sind. Abzielend auf die Beobachtung von Heterogenitäten im G4 RNS Faltungsweg oder

Multimerisierungsschritten, wurden Einzelmolekül (sm) Experimente unter Verwendung von Försterenergietransfer (FRET) durchgeführt.

Kapitel 4 behandelt folglich die Anwendung von Einzelmolekül-FRET an G4 RNS. Ein Arbeitskonstrukt für den Einsatz bei FRET Studien wird hergestellt in dem eine kommerzielle NRAS Sequenz eigenhändig mit einem SCy5 Fluorophor am 5'-Ende mittels NHS-Ester-Kopplung in gepufferten wässrigen Bedingungen markiert wird. Das erfolgreiche Markieren wird durch Massenspektrometrie der SCy5 NRAS RNS gezeigt. Durch Verwendung eines kommerziellen DNS Oligonukleotids, intern markiert mit SCy3 und mit Biotin am 5'-Ende, werden die RNS:DNS Hybridisierungsbedingungen optimiert. Immobilisierung des Konstrukts auf Quartzobjektträgern wird danach mittels Biotin Streptavidin-Verknüpfung erreicht und einem Arbeitssystem für Einzelmolekülbeobachtung der NRAS G4 aufgesetzt.

Einzelmolekül FRET Filme werden in der Anwesenheit und Abwesenheit von K^+ (0 – 100 mM) unter Verwendung von alternierender Laseranregung (ALEX) aufgenommen, um diese Moleküle auszuwählen welche beide Markierungen tragen. Antikorrelation von Donor- mit Akzeptorfluoreszenz wird beim Ausbleichen von SCy5 beobachtet (*Abb. 8.3.*), was anzeigt, dass FRET zwischen den Donor- und Akzeptorfarbstoffen auftritt. Trotzdem sind keine Übergänge zwischen den FRET Zuständen vorhanden und nur einen höches FRET (~ 1) wird beobachtet. Wahrscheinlich bleibt die NRAS G4 in einer stabilen gefalteten Konformation ohne dass Faltungsintermediate auftreten. Dies ist nicht überraschend, wenn man die hohe NRAS RNS Stabilität berücksichtigt, die in Kapitel 2 beschrieben wurde. Im Gegensatz durch, tritt keine NRAS Multimerisierung auf der Einzelmolekülebene auf. Zum einen verhindert die geringen Konzentration im pM Bereich und die Anwesenheit eines raumfordernden SCy5 Fluorophores am 5'-Ende der RNS die Aggregation während der Präparation, und zum anderen verhindert die Oberflächenbindung des einzeln Molekühls die Interaktion und damit die Multimerisierung nachdem überflüssiges Moleküle weggewaschen worden.

Die gewählte Strategie für die Einzelmolekülbeobachtung der NRAS G4 ergibt eine stabile G4 gefaltete Konformation mit einem hohen FRET Wert. Es werden keine Dynamiken beobachtet, sogar in Abwesenheit von K^+ Kationen im Bildgebungspuffer. Wir schlagen deshalb eine Strategie unter Verwendung des G4 spezifischen G4 Resolvase Enzyms, um die G4 Entfaltung in Echtzeit zu messen. Dieses Schema würde Vergleichsexperimente mit stabilisierenden Liganden erlauben und könnte ein deutliches Bild über die Zeitskala der G4 Dynamik unter physiologischen Bedingungen, die G-Quadruplexbildung und deren Ligandwechselwirkung *in vivo* ergeben.

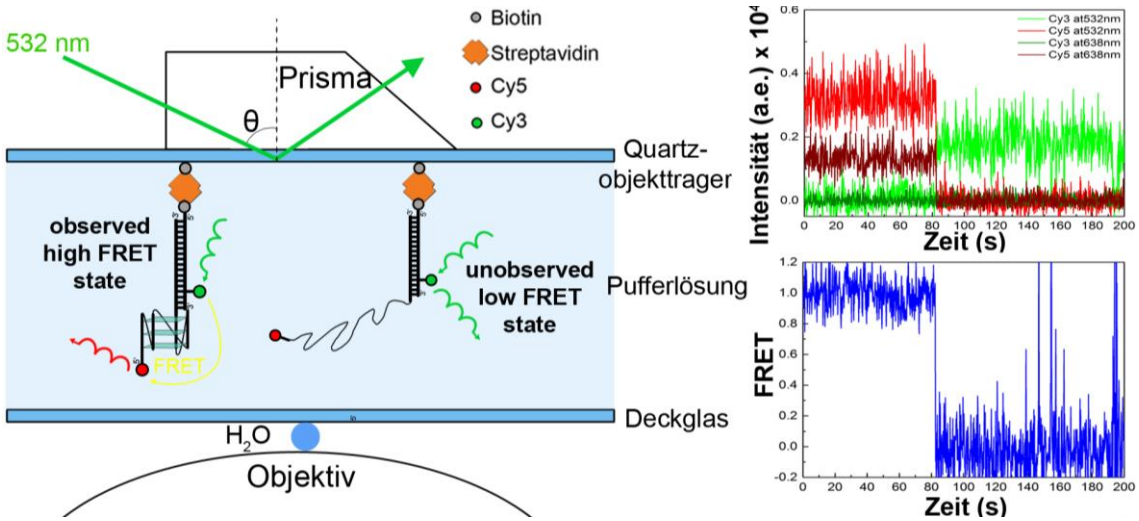


Abbildung 8.3. Einzelmolekül FRET Messungen der NRAS G-Quadruplex. Nur der hohe FRET Zustand, der zu einer gefalteten und stabilen G4 gehört, wird beobachtet. Ein Beispiel einer typische Zeitlinie wird zusammen mit dem jeweiligen auftretenden FRET Wert gezeigt. In diesem Molekül, gemessen ohne K^+ Zugabe, tritt Bleichen von SCy5 bei 80 s auf und Antikorrelation zwischen Donor- und Akzeptorfluoreszenz wird beobachtet. Dies bestätigt, dass FRET zwischen beiden Farbstoffen auftritt.

Übergreifend betrachtet diese Arbeit das *in vitro* Verhalten von zwei kurzen G-reichen RNS-Sequenzen, die dafür bekannt sind, in G-Quadruplex Motive zu falten. Ihre Metallionenabhängigkeit wird zusammen mit ihrer Tendenz gegenüber Multimerisierung untersucht und beide Aspekte werden mit dem existierenden Wissen über G4 DNS verglichen. Auf diese Weise wird eine gute Grundlage für die besten Pufferbedingungen und die Wahl der monomeren RNS Sequenzen für *in vitro* Experimente etabliert. Die Beobachtungen von NRAS G4 auf dem Einzelmoleküllevel unterstreichen weiter, dass diese G4 RNS sehr stabil sind, so sehr, dass keine Dynamiken, weder in Anwesenheit noch Abwesenheit von K^+, innerhalb des Zeitrahmens der Experimente gemessen werden. Die hier gezeigten Ergebnisse repräsentieren einen weiteren Schritt Richtung Verständnis dieser RNS Motive, welche für die *in vivo* Regulierung von biologischen Prozessen relevant sind und Potential als zukünftige Antitumorziele haben.

9. Resumen

Las secuencias de ácidos nucleicos ricas en guaninas tienen tendencia a organizarse en estructuras helicoidales conocidas como G-cuadruplejos (G4). Estos consisten en un apilamiento de dos o más G-cuartetos, cada uno de los cuales está formado por una disposición planocuadrada de bases guaninas enlazadas por puentes de hidrógeno. Los G4 se encuentran en regiones reguladoras del transcriptoma, especialmente en regiones sin traducir del extremo 5' (5' UTR), y se pueden detectar en el citoplasma de células humanas immobilizadas. Tienen una alta estabilidad termodinámica en condiciones fisiológicas y se pliegan siempre de forma paralela.

Nos interesamos por los ARN en G4 en tanto que elementos reguladores presentes *in vivo*, así como dianas potenciales para nuevas terapias: por ejemplo, terapias antitumorales. El grupo hidroxilo extra, 2'-OH, presente solo en ARN, permite el diseño y desarrollo de pequeñas moléculas capaces de actuar selectivamente sobre G4s de ARN, diferenciándolos de los de ADN. Aquí trabajamos con la secuencia de 18 nucleótidos 5'-GGGAGGGGCGGGUCUGGG-3', del extremo 5' sin traducir del ARN mensajero (mRNA) del oncógeno NRAS (neuroblastoma rat sarcoma viral oncogene homolog), que codifica una proteína relacionada con la proliferación celular. Se ha demostrado previamente que cuando NRAS se pliega en forma de G4 es capaz de reducir la síntesis de proteína *in vitro* mediante inhibición de la traducción. Además, se ha visto que un ligando puede actuar sobre el cuadruplejo NRAS, que se perfila, así, como una diana interesante para nuevos esfuerzos antitumorales. También utilizamos el ARN que contiene la repetición telomérica (TERRA) y que es el producto de transcripción de la repetición telomérica de ADN, siendo ampliamente estudiado. Gracias a las aportaciones de otros autores se sabe que la secuencia de 24 nucleótidos 5'-$(UUAGGG)_4$-3' se pliega en forma de G4 en presencia de K^+ e inhibe la actividad de la telomerasa. Esta enzima se sobreexpresa en aproximadamente 80 % de los cánceres humanos y es responsable de la inmortalidad de las células tumorales. Por todo lo expuesto, TERRA, así como su análogo de ADN, está siendo investigado desde el punto de vista de dianas tumorales, ya que el cuadruplejo TERRA estabilizado *in vivo* tendría un efecto anticancerígeno por vía de la inhibición de la telomerasa.

Como todos los ADN y ARN, y debido a su esqueleto de fosfatos, los G-cuadruplejos tienen una alta carga negativa y, por lo tanto, son sensibles a las condiciones catiónicas circundantes. Además, los grupos carbonilos de las guaninas que forman los G-cuartetos contienen oxígenos con pares libres de electrones, que originan una carga negativa localmente alta en el centro de los planos. De este modo, se forma un lugar específico para el enlace de cationes, capaz de coordinar iones metálicos, necesarios para el plegamiento y la estabilidad de las estructuras en G4 (*Fig. 9.1.*). Como consecuencia de este lugar de enlace

definido, se predice que solamente los cationes con un tamaño, carga y energía de deshidratación adecuados interaccionarán con los G4s.

Varias publicaciones tratan la interacción entre G4 de ADN y iones metálicos. En cambio, hay menos información disponible sobre G4 de ARN. Se sabe que estos son más estables que sus análogos de ADN aunque no se han referido estudios sistemáticos con iones metálicos. Por consiguiente, en este trabajo tenemos como objetivo entender las interacciones entre iones metálicos y G4 de ARN, así como el efecto de los cationes en su plegamiento y estabilidad, comparando los resultados con los datos conocidos para ADN. Asimismo, queremos observar la dinámica dependiente de iones metálicos al nivel de moléculas únicas para elucidar las etapas del plegamiento del ARN en G4.

El Capítulo 2 tiene como objetivo de dilucidar el efecto de iones metálicos mono (Li^+, Na^+, K^+, Rb^+, Cs^+, y NH_4^+) y divalentes (Mg^{2+}, Ca^{2+}, Sr^{2+} y Ba^{2+}) en el plegamiento y la estabilidad de ARN que forman G4. Utilizamos dicroísmo circular (CD, de las siglas en inglés) y espectros UV de diferencia térmica para verificar la formación de G-cuadruplejos con NRAS y TERRA que, se comprueba con las señales de los espectros obtenidos, forman G4s paralelos. A continuación, se llevan a cabo experimentos UV y de dicroísmo circular para determinar el punto de fusión y evaluar el grado de plegamiento y la estabilidad relativa, respectivamente, en cada una de las condiciones iónicas examinadas. Los resultados de dicroísmo circular en agua revelan que los G4s se encuentran preplegados sin adición de cationes, con una ligera señal de CD originada por el preapilamiento de las guaninas. La subsecuente adición de iones metálicos mono o divalentes aumenta la forma plegada en G4, evidenciada con un aumento en intensidad de las señales de CD. En particular, Na^+, K^+, Sr^{2+}, y Rb^+ promueven el plegamiento en G4 de NRAS y TERRA, mientras que Ba^{2+} y NH_4^+ favorecen solo el plegamiento de la secuencia NRAS. En una concentración más elevada, Ca^{2+} aumenta también el plegamiento en G4 para las dos secuencias de ARN, y Mg^{2+}, el de NRAS.

Al mismo tiempo, solo K^+ y Sr^{2+} resultan en un aumento significativo de las temperaturas de fusión de NRAS y TERRA, que son ya moderadamente estables en agua sola (T_m's ~ 50 °C). Esta mayor estabilización de los G4 en presencia de K^+ y Sr^{2+} ya es conocida para el ADN y se explica por el hecho de que los dos cationes tienen un tamaño y energía de deshidratación óptimos para ajustarse entre los G-cuartetos, lo que forma un canal de cationes. A diferencia del ADN, en los dos ARN no hay ninguna estabilización observada con la adición de Na^+, incluso a altas concentraciones de hasta 100 mM. Este hecho es probablemente debido a la preferencia de los G4 paralelos por enlazar K^+ y también a que los cuadruplejos de ARN son siempre paralelos, mientras que los ADN en G4 son más heterogéneos y pueden plegarse también de forma antiparalela o híbrida, dependiendo de las condiciones circundantes.

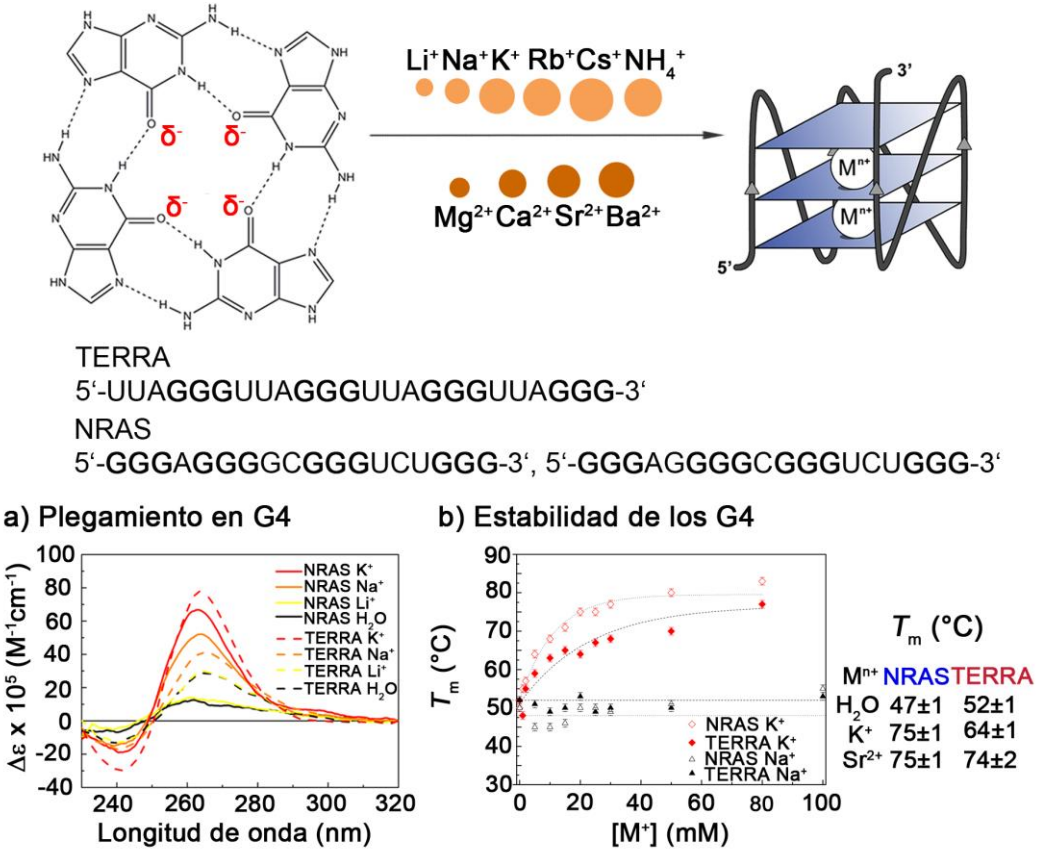


Figura 9.1. Interacción de ARNs en G4 (TERRA y NRAS) con iones metálicos mono y divalentes (Li^+, Na^+, K^+, Rb^+, Cs^+, y NH_4^+; y Mg^{2+}, Ca^{2+}, Sr^{2+}, y Ba^{2+}). NRAS puede formar dos motivos diferentes en G4, dependiendo de qué G están presentes en los G-cuartetos. a) El espectro de CD en agua muestra que Na^+ y K^+, entre otros (*no mostrados aquí*), promueven el plegamiento en G4; b) El K^+ aumenta la estabilidad de los G4 de NRAS y TERRA, mientras que el Na^+ no tiene ningún efecto. Efectivamente, K^+ y Sr^{2+} son los dos únicos cationes que resultan en un aumento significativo de las temperaturas de fusión (T_m) de los ARNs, comparado con los controles medidos en agua.

Hemos descrito la dependencia catiónica de dos ARN que pliegan en G4 paralelos y la hemos comparado con la interacción catiónica conocida para ADN. La diferencia principal es la insensibilidad al Na^+, que demuestra la estabilidad de los ARN; así como el hecho que los ARN en G4 son ya bastante estables en agua a diferencia de sus homólogos de ADN. Además, hemos observado algunas diferencias en la interacción con cationes entre las dos secuencias: el ARN NRAS es más sensible a la adición de cationes, como se ve con la influencia de Ba^{2+}, NH_4^+ y Mg^{2+} en su plegamiento. El hecho que estos cationes tengan un efecto significativo en NRAS y no efecto mesurable en TERRA da a pensar que diferencias en la respuesta a cationes podrían ser potencialmente explotadas para ejercer una acción específica hacia un ARN en G4 concreto. En definitiva, el estudio sistemático de la dependencia catiónica de ARN en G4 enfatiza la importancia de esclarecer las condiciones de trabajo óptimas para este tipo de secuencias, lo que permite a los investigadores el control del plegamiento en G4 y de la estabilidad de los motivos resultantes.

Los estudios espectroscópicos descritos en el Capítulo 2 dan algunas indicaciones de posible multimerización o agregación, especialmente para el caso de NRAS. Por ejemplo, las temperaturas de fusión no son independientes de la concentración de NRAS por encima de 10-15 µM, lo que indica que, aparte del G4 intramolecular, hay posiblemente multímeros presentes en disolución. La multimerización de G4 ya ha sido reseñada en la bibliografía para secuencias cortas de ADN y ARN, especialmente conteniendo *loops*‡ cortos o guaninas consecutivas en los extremos.

El Capítulo 3 trata, así pues, de la tendencia a la multimerización del cuadruplejo NRAS en un exceso de cationes estabilizantes. Esta secuencia de ARN forma una unidad paralela de G4 que multimeriza fácilmente en presencia de K^+, Sr^{2+} o NH_4^+. La electroforesis con geles de poliacrilamida (PAGE) en condiciones nativas y en presencia de los cationes mencionados muestra una banda de movilidad reducida, que podría corresponder a un dímero de NRAS. Este dímero de G4 se confirma por espectrometría de masas en condiciones de ionización suaves (ESI-MS). Por el contrario, TERRA aparece como un monómero de G4 tanto en PAGE como en MS. La determinación de los radios hidrodinámicos (r_H) en soluciones de K^+ o Sr^{2+} se lleva a cabo vía dispersión de luz dinámica (DLS) y confirma la aparición de estados de más alta multimerización a concentraciones de cationes mono o divalentes más elevadas. Además, los valores de r_H para NRAS se corresponden otra vez con una mayor tendencia a la multimerización para esta secuencia.

Con el fin de entender la diferencia en la tendencia a la multimerización entre NRAS y TERRA, se adicionaron a NRAS nucleótidos flanqueantes de la secuencia natural del mRNA y se volvió a estudiar con PAGE y ESI-MS la tendencia a multimerizar de las secuencias modificadas. La adición de un nucleótido a cada extremo de la secuencia (NRAS20) resulta en dimerización. Por el contrario, dos bases adicionales a cada extremo, 5’ y 3’ (NRAS22), impiden la multimerización, lo que resulta en un monómero de NRAS. Además, ya no se agrega una secuencia de NRAS con “UUA” adicionado al extremo 5’ (NRAS21). En paralelo, al eliminar los tres primeros nucleótidos (UUA) de la secuencia TERRA se obtiene una secuencia más corta (TERRA21) que dimeriza. Estos resultados de PAGE sugieren que la razón de la multimerización de NRAS18 es la presencia de tres guaninas consecutivas en el extremo 5’, y que el dímero se forma presumiblemente por vía de apilamiento 5’-5’. Los experimentos de MS corroboran que el dímero de NRAS contiene seis cuartetos y cinco iones metálicos (*Fig. 9.2.*).

‡ Se conocen como *loops* los nucleótidos que conectan los cuartetos de guaninas.

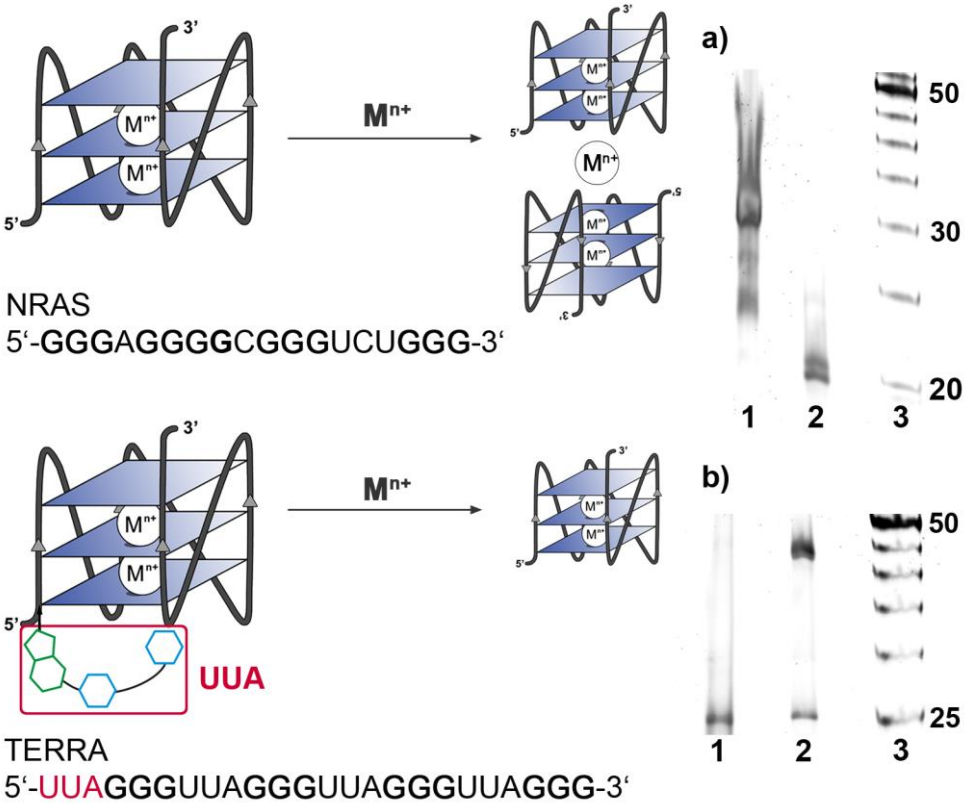


Figura 9.2. Comportamiento de multimerización de los ARN en G4 (NRAS y TERRA) en exceso de cationes estabilizantes (K^+, NH_4^+ y Sr^{2+}). Se observa el efecto de las secuencias flanqueantes, que resulta en un TERRA monomérico y un NRAS dimérico. a) PAGE nativo en 20 mM K^+ mostrando NRAS18, dímero (1), y NRAS21, monómero (2); b) PAGE nativo en 20 mM K^+ mostrando TERRA24, monómero (1), y TERRA21, una mezcla de monómero y dímero (2).

En el presente trabajo confirmamos que los ARN en G4 presentan una mayor tendencia hacia la multimerización comparados con los ADN en G4. Esto es probablemente debido a que la forma paralela permite un mejor apilamiento de tipo π-π entre los monómeros. El cuadruplejo NRAS multimeriza en forma de dímero en K^+, Sr^{2+} o NH_4^+; mientras que TERRA es un monómero intramolecular en las mismas condiciones (*Fig. 9.2.*). Hemos logrado explicar esta diferencia entre los dos ARN por medio de experimentos de PAGE y MS con secuencias modificadas que muestran que la multimerización está dirigida por el extremo 5' y la presencia ahí de guaninas consecutivas formando G4. Se tendrán que tomar en cuenta estos resultados para prevenir la agregación durante el diseño de nuevas secuencias cortas de ARN en G4 para experimentos *in vitro*.

Además del G4 plegado y de la conformación abierta, pueden encontrarse otras formas en disolución si el ARN en G4 multimeriza, resultando en la observación de un promedio heterogéneo en los experimentos espectroscópicos y en los geles descritos en los Capítulos 2 y 3. Este hecho, junto con intermediarios potenciales que aparecerían durante el plegamiento por adición de cationes, resultaría en heterogeneidades, que no son observables

en las medidas de conjunto descritas hasta ahora. Así pues, con el objetivo de observar las heterogeneidades en la ruta de plegamiento o multimerización de ARN en G4, se llevaron a cabo experimentos en condiciones de molécula única (*single molecule, sm*) utilizando la técnica de transferencia de energía de resonancia Förster (FRET).

El Capítulo 4 explora la aplicación de smFRET al estudio de un RNA en G4. El constructo para medidas de FRET se diseñó con una secuencia NRAS comercial, que se marca con un fluoróforo SCy5 al extremo 5', por medio de un acoplamiento NHS-ester en condiciones acuosas y bajo control del pH. El marcado exitoso se demuestra vía MS del ARN SCy5-NRAS. Utilizando un oligonucleótido de ADN comercial, marcado internamente con SCy3 y conteniendo biotina al extremo 5', se optimizan las condiciones de hibridación ARN:ADN. La inmovilización del constructo resultante en un portaobjetos de cuarzo se consigue vía enlace de la biotina con la estreptadivina, y así se establece un sistema de trabajo para la observación del cuadruplejo NRAS en experimentos de molécula única.

Se registran películas de smFRET en presencia y ausencia de K^+ (0-100 mM), utilizando excitación láser alterna (ALEX) para seleccionar las moléculas que contienen ambos fluoróforos. Anticorrelación de la fluorescencia del donador con el aceptor es observada al mismo tiempo que el blanqueo de SCy5 (*Fig. 9.3.*), lo cual indica que FRET ocurre entre el donador y el aceptor. Sin embargo, no hay transiciones presentes entre estados de FRET y solo se observa un alto estado de FRET (~ 1). Presumiblemente, el NRAS en G4 permanece en una conformación plegada y estable y no se observan intermediarios. Esto no es sorprendente, considerando la alta estabilidad del ARN NRAS observada en el Capítulo 2. Al mismo tiempo, en condiciones de molécula única no se observa multimerización de NRAS. Por un lado, la baja concentración en el rango de pM y a la presencia de un fluoróforo SCy5 voluminoso al extremo 5' del ARN dificultan la agregación durante la preparación de las muestras, y por el otro, la immobilización en superficie impide la aparición de interacciones ARN-ARN una vez el NRAS en exceso ha sido eliminado por medio de un lavado con solución tampón.

La estrategia escogida para la observación del cuadruplejo NRAS en experimentos de molécula única resulta en una conformación en G4 plegada y estable con un alto valor de FRET. No se observa ninguna dinámica ni estados intermedios de FRET, incluso sin catión K^+ presente en la disolución. Así pues, proponemos cómo perspectiva una estrategia que utilizaría la enzima G4 resolvasa, específica para G4, para desenrollar la estructura, permitiendo medir su despliegue en tiempo real. Este esquema permitiría experimentos de competición con ligandos estabilizantes y podría dar una idea clara de si la dinámica de G4 en condiciones fisiológicas podría o no permitir la formación de G4 y su interacción con ligandos *in vivo*.

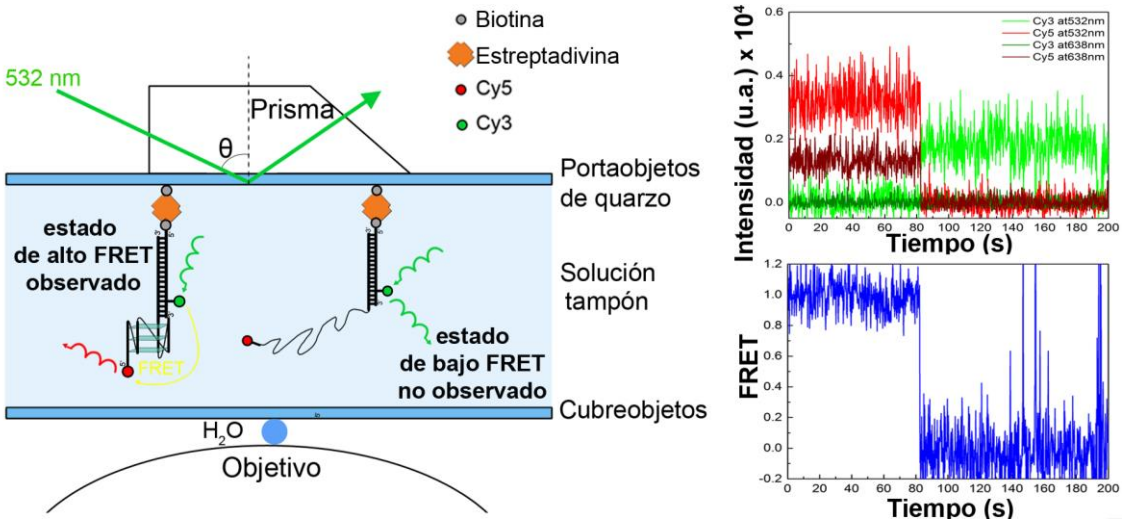


Figura 9.3. Medidas de smFRET del cuadruplejo NRAS. Solamente se observa el estado de alto FRET, correspondiente a un G4 plegado y estable. Aquí se muestra un ejemplo típico de traza temporal, junto con el respectivo valor de FRET aparente. En esta molécula, medida sin adición de K^+, el blanqueo de SCy5 ocurre a los 80 s, y se observa anticorrelación entre la fluorescencia del donador y del aceptor. Esto confirma que FRET ocurre entre los dos fluoróforos.

En conjunto, esta tesis se centra en el comportamiento *in vitro* de dos secuencias cortas de ARN ricas en guaninas y conocidas por formar G-cuadruplejos. Se describe su dependecia de iones metálicos, junto con su tendencia a la multimerización, y estos dos aspectos son comparados con el conocimiento existente sobre ADN en G4. De este modo, se establece una buena base para las condiciones de tampón óptimas y la elección de secuencias de ARN monoméricas para experimentos *in vitro*. Las observaciones del cuadruplejo NRAS en experimentos de molécula única enfatizan la gran estabilidad de estos ARN en G4, tan estables que no se detecta ninguna dinámica ni en presencia ni en ausencia de K^+ durante el periodo de tiempo de los experimentos. Los resultados descritos aquí representan otro paso hacia el conocimiento de estas estructuras de ARN, que son relevantes para la regulación de procesos biológicos *in vivo* y tienen potencial como futuras dianas antitumorales.

10. Bibliography

1. Levene, P. A. “The structure of yeast nucleic acid. III. Ammonia hydrolysis”. *J. Biol. Chem.* 33, 425-428 (1918).
2. Levene, P. A. “The structure of yeast nucleic acid”. *Studies from the Rockefeller Institute for Medical Research* 30, 221 (1919).
3. Crick, F. H. C. “On protein synthesis”. *Sym. Soc. Exp. Biol.* 12, 138-163 (1958).
4. Geiduschek, E. P.; Haselkorn, R. “Messenger RNA”. *Annu. Rev. Biochem.* 38, 647-676 (1969).
5. Editorial. “A windfall for RNA. Two 2006 Nobel prizes reflect the central role of RNA in gene regulation and emphasize the interplay of discoveries in chemistry and biology”. *Nat. Chem. Biol.* 3, 1 (2007).
6. Zamore, P. D. “RNA interference: big applause for silencing in Stockholm”. *Cell* 127, 1083-1086 (2006).
7. DeVincenzo, J. P. “Harnessing RNA interference to develop neonatal therapies: from Nobel Prize winning discovery to proof of concept clinical trials”. *Early Hum. Dev.* 85, S31-5 (2009).
8. Pérez-Cañadillas, J.-M.; Varani, G. “Recent advances in RNA–protein recognition”. *Curr. Opin. Struc. Biol.* 11, 53-58 (2001).
9. Sullenger, B. A.; Gilboa, E. “Emerging clinical applications of RNA”. *Nature* 418, 252-258 (2002).
10. Puerta-Fernández, E.; Romero-López, C.; Barroso-delJesus, A.; Berzal-Herranz, A. “Ribozymes: recent advances in the development of RNA tools”. *FEMS Microbiol. Rev.* 27, 75-97 (2003).
11. Guo, P. “RNA Nanotechnology: Engineering, Assembly and Applications in Detection, Gene Delivery and Therapy”. *J. Nanosci. Nanotechno.* 5, 1964-1982 (2005).
12. Filbin, M. E.; Kieft, J. S. “Toward a structural understanding of IRES RNA function”. *Curr. Opin. Struc. Biol.* 19, 267-276 (2009).
13. Guan, L.; Disney, M. D. “Recent advances in developing small molecules targeting RNA”. *ACS Chem. Biol.* 7, 73-86 (2012).
14. Ozsolak, F.; Milos, P. M. “RNA sequencing: advances, challenges and opportunities”. *Nat. Rev. Genet.* 12, 87-98 (2011).
15. Rajkowitsch, L.; Chen, D.; Stampfl, S.; Semrad, K.; Waldsich, C.; Mayer, O.; Jantsch, M. F.; Konrat, R.; Bläsi, U.; Schroeder, R. “RNA Chaperones, RNA Annealers and RNA Helicases”. *RNA Biol.* 4, 118-130 (2014).
16. Gebauer, F.; Hentze, M. W. “Molecular mechanisms of translational control”. *Nat. Rev. Mol. Cell Bio.* 5, 827-835 (2004).
17. Doudna, J. A. “Structural genomics of RNA”. *Nat. Struct. Biol.,* 954-956 (2000).
18. Williamson, J. R. “Induced fit in RNA-protein recognition”. *Nat. Struct. Biol.* 7, 834-837 (2000).
19. Jackson, R. J.; Hellen, C. U. T.; Pestova, T. V. “The mechanism of eukaryotic translation initiation and principles of its regulation”. *Nat. Rev. Mol. Cell Bio.* 11, 113-127 (2010).
20. Noller, H. F. “Ribosomal RNA and Translation”. *Annu. Rev. Biochem.* 60, 191-227 (1991).
21. Quigley, G.; Rich, A. “Structural domains of transfer RNA molecules”. *Science* 194, 796-806 (1976).

22. Lilley, D. M. “The origins of RNA catalysis in ribozymes”. *Trends Biochem. Sci.* 28, 495-501 (2003).

23. Cech, T. R. “The Chemistry of Self-Splicing RNA and RNA Enzymes”. *Science* 236, 1532-1539 (1987).

24. Steitz, T. A.; Moore, P. B. “RNA, the first macromolecular catalyst: the ribosome is a ribozyme”. *Trends Biochem. Sci.* 28, 411-418 (2003).

25. Kruger, K.; Grabowski, P. J.; Zaug, A. J. ; Sands, J. ; Gottschling, D. E.; Cech, T.R. *“*Self-splicing RNA: Autoexcision and autocyclization of the ribosomal RNA intervening sequence of tetrahymena”. *Cell* 31, 147-157 (1982).

26. Guerrier-Takada, C.; Gardiner, K.; Marsh, T. Pace, N.; Altman, S. “The RNA Moiety of Ribonuclease P Is the Catalytic Subunit of the Enzyme”. *Cell* 35, 849-857 (1983).

27. Serganov, A.; Patel, D. J. “Ribozymes, riboswitches and beyond: regulation of gene expression without proteins”. *Nat. Rev. Genet.* 8, 776-790 (2007).

28. Copley, S. D. Smith, E.; Morowitz, H. J. “The origin of the RNA world: co-evolution of genes and metabolism”. *Bioorg. Chem.* 35, 430-443 (2007).

29. Neveu, M. Kim, H.-J.; Benner, S. A. “The "strong" RNA world hypothesis: fifty years old”. *Astrobiology* 13, 391-403 (2013).

30. Gilbert, W. “The RNA world”. *Nature* 319, 618 (1986).

31. Cech, T. R. “The RNA worlds in context”. *CSH Perspect. Biol.* 4, a006742 (2012).

32. Schrum, J. P.; Zhu, T. F.; Szostak, J. W. “The origins of cellular life”. *CSH Perspect. Biol.* 2, a002212 (2010).

33. Powner, M. W.; Gerland, B.; Sutherland, J. D. “Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions”. *Nature* 459, 239-242 (2009).

34. Marlaire, R. *NASA Ames Reproduces the Building Blocks of Life in Laboratory* (2015).

35. Robertson, M. P.; Joyce, G. F. “The origins of the RNA world”. *CSH Perspect. Biol.* 4 (2012).

36. Larralde, R.; Robertson, M. P.; Miller, S. L. “Rates of decomposition of ribose and other sugars: Implications for chemical evolution”. *P. Natl. Acad. Sci. USA* **92,** 8158-8160 (1995).

37. Joyce, G. F.; Orgel, L. E. “Prospects for Understanding the Origin of the RNA World”. In *The RNA World*. 3rd edition, ed. Gesteland, R. F. (Cold Spring Harbor Laboratory Press, 2005).

38. Schmidt, J. G.; Nielsen, P. E.; Orgel, L. E. “Information transfer from peptide nucleic acids to RNA by template-directed syntheses”. *Nucleic Acids Res.* 25, 4797-4802 (1997).

39. Joyce, G. F. “The antiquity of RNA-based evolution”. *Nature* 418, 214-221 (2002).

40. Shi, H.; Moore, P. B. “The crystal structure of yeast phenylalanine tRNA at 1.93 Å resolution: A classic structure revisited”. *RNA* 6, 1091-1105 (2000).

41. Connelly, S.; Manley, J. L. “A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II”. *Genes Dev.* 2, 440-452 (1988).

42. Colgan, D. F.; Manlex, J. L. “Mechanism and regulation of mRNA polyadenylation”. *Genes Dev.* 11, 2755-2766 (1997).

43. Jiao, X.; Chang, J. H.; Kilic, T.; Tong, L.; Kiledjian, M. “A mammalian pre-mRNA 5' end capping quality control mechanism and an unexpected link of capping to pre-mRNA processing”. *Mol. Cell* 50, 104-115 (2013).

44. Newmann, A. J. “Pre-mRNA splicing”. *Curr. Opin. Gen. Dev.* 4, 298-304 (1994).

45. Faustino, N. A.; Cooper, T. A. “Pre-mRNA splicing and human disease”. *Genes Dev.* 17, 419-437 (2003).

46. Cheng, H.; Dufu, K.; Lee, C.-S.; Hsu, J. L.; Dias, A.; Reed, R. *“*Human mRNA export machinery recruited to the 5' end of mRNA”. *Cell* 127, 1389-1400 (2006).

47. Hansen, J. L.; Long, A. M.; Schultz, S. C. “Structure of the RNA-dependent RNA polymerase of poliovirus”. *Structure* 5, 1109-1122 (1997).

48. Goff, S. P. “Retroviral Reverse Transcriptase: Synthesis, Structure, and Function”. *J. A. I. D. S.* 3, 817-831 (1990).

49. Cordaux, R.; Batzer, M. A. “The impact of retrotransposons on human genome evolution”. *Nat. Rev. Genet.* 10, 691-703 (2009).

50. Linger, J.; Hughes, T. R.; Shevchenko, A.; Mann, M.; Lundblad, V.; Cech, T. R. “Reverse Transcriptase Motifs in the Catalytic Subunit of Telomerase”. *Science* 276, 561-567 (1997).

51. Usman, N.; Cedergren, R. “Exploiting the chemical synthesis of RNA”. *Trends Biochem. Sci.* 17, 334-339 (1992).

52. Milligan, J. F.; Uhlenbeck, O. C. “Synthesis of small RNAs using T7 RNA polymerase”. In *RNA Processing Part A: General Methods* (Elsevier 1989).

53. Mattick, J. S. “The Functional Genomics of Noncoding RNA”. *Science* 309, 1527-1528 (2005).

54. Taft, R. J.; Pheasant, M.; Mattick, J. S. “The relationship between non-protein-coding DNA and eukaryotic complexity”. *BioEssays* 29, 288-299 (2007).

55. Costa, F. F. “Non-coding RNAs, epigenetics and complexity”. *Gene* 410, 9-17 (2008).

56. Lodish, H.; Berk, A.; Zipursky, L. S.; Matsudaira, P.; Baltimore, D.; Darnell, J. *Molecular Cell Biology* (W. H. Freeman & Co Ltd, New York, USA, 2000).

57. Eddy, S. R. “Non-coding RNA genes and the modern RNA world”. *Nat. Rev. Genet.* 2, 919-929 (2001).

58. Castanotto, D.; Rossi, J. J. “The promises and pitfalls of RNA-interference-based therapeutics”. *Nature* 457, 426-433 (2009).

59. Henkin, T. M. “Riboswitch RNAs: using RNA to sense cellular metabolism”. *Genes Dev.* 22, 3383-3390 (2008).

60. Henkin, T. M. “Riboswitch RNAs: using RNA to sense cellular metabolism”. *Genes Dev.* 22, 3383-3390 (2008).

61. Zhang, J.; Lau, M. W.; Ferré-D'Amaré, A. R. “Ribozymes and riboswitches: modulation of RNA function by small molecules”. *Biochemistry* 49, 9123-9131 (2010).

62. Drew, H. R.; Wing, R. M.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. *“*Structure of a B-DNA dodecamer: Conformation and dynamics”. *P. Natl. Acad. Sci. USA* 78, 2179-2183 (1981).

63. Dock-Bregeon, A. C.; Chevrier, B.; Podjarny, A.; Johnson, J.; de Bear, J. S.; Gough, G. R.; Gilham, P. T.; Moras, D. *“*Crystallographic structure of an RNA helix: $[U(UA)_6A]_2$”. *J. Mol. Biol.* 209, 459-474 (1989).

64. Neidle, S. *Principles of nucleic acid structure.* 1st ed. (Elsevier; Academic Press, Amsterdam, Boston, 2008).

65. Sinden, R. R. *DNA Structure and Function* (Academic Press, San Diego, California, USA, 1994).

66. Drancourt, M.; Aboudharam, G.; Signoli, M.; Dutour, O.; Raoult, D. “Detection of 400-year-old Yersinia pestis DNA in human dental pulp: An approach to the diagnosis of ancient septicemia”. *P. Natl. Acad. Sci. USA* 95, 12637-12640 (1998).

67. Pääbo, S.; Poinar, H.; Serre, D.; Jaenicje-Despres, V.; Hebler, J.; Rohland, N.; Kuch, M.; Krause, J.; Vigilant, L.; Hofreiter, M. “Genetic analyses from ancient DNA”. *Annu. Rev. Genet.* 38, 645-679 (2004).

68. Papagrigorakis, M. J.; Yapijakis, C.; Synodinos, P. N.; Baziotopoulou-Valavani, E. “DNA examination of ancient dental pulp incriminates typhoid fever as a probable cause of the Plague of Athens”. *Int. J. Infect. Dis.* 10, 206-214 (2006).

69. Leontis, N. B.; Westhof, E. “Geometric nomenclature and classification of RNA base pairs”. *RNA* 7, 499-512 (2001).

70. Parkinson, G. N. “Fundamentals of Quadruplex Structures”. *Quadruplex Nucleic Acids,* 1-30 (2006).

71. Draper, D. E.; Grilley, D.; Soto, A. M. “Ions and RNA folding”. *Annu. Rev. Bioph. Biom.* 34, 221-243 (2005).

72. Ramesh, A.; Winkler, W. C. “Magnesium-sensing riboswitches in bacteria”. *RNA Biol.* 7, 77-83 (2014).

73. DeRose, V. J. “Metal ion binding to catalytic RNA molecules”. *Curr. Opin. Struc. Biol.* 13, 317-324 (2003).

74. Shapiro, B. A.; Yingling, Y. G.; Kasprzak, W.; Bindewald, E. “Bridging the gap in RNA structure prediction”. *Curr. Opin. Struc. Biol.* 17, 157-165 (2007).

75. Reyes, F. E.; Garst, A. D.; Batey, R. T. “Strategies in RNA Crystallography”. In *Biophysical, Chemical, and Functional Probes of RNA Structure, Interactions and Folding: Part B* (Elsevier 2009).

76. Varani, G.; Aboul-ela, F.; Allain, F. H.-T. “NMR investigation of RNA structure”. *Prog. Nucl. Mag. Res. Sp.* 29, 51-127 (1996).

77. Alberts, B.; Wilson, J.; Hunt, T. *Molecular biology of the cell.* 5th ed. (Garland Science, New York, 2008).

78. Lorsch, J. R.; Szostak, J. W. “In Vitro Selection of RNA Aptamers Specific for Cyanocobalamin”. *Biochemistry* 33, 973-982 (1994).

79. Kisseleva, N.; Khvorova, A.; Westhof, E.; Schiemann, O. “Binding of manganese(II) to a tertiary stabilized hammerhead ribozyme as studied by electron paramagnetic resonance spectroscopy”. *RNA* 11, 1-6 (2005).

80. Schnabl, J.; Sigel, R. K O. “Controlling ribozyme activity by metal ions”. *Curr. Opin. Chem. Biol.* 14, 269-275 (2010).

81. Gallo, S.; Oberhuber, M.; Sigel, R. K. O.; Kräutler, B. “The corrin moiety of coenzyme B12 is the determinant for switching the btuB riboswitch of E. coli”. *Chembiochem* 9, 1408-1414 (2008).

82. Athavale, S. S.; Petrov, A. S.; Hsiao, C.; Watkins, D.; Prickett, C. D.; Gossett, J. J.; Lie, L.; Bowman, J. C.; O’Neill, E.; Bernier, C. R.; Hud, N. V.; Wartell, R. M.; Harvey, S. C.; Williams, L. D. “RNA Folding and Catalysis Mediated by Iron (II)”. *PLoS ONE* 7, e38024 (2012).

83. Draper, D. E. “RNA folding: thermodynamic and molecular descriptions of the roles of ions”. *Biophys. J.* 95, 5489-5495 (2008).

84. Draper, D. E. “A guide to ions and RNA structure”. *RNA* 10, 335-343 (2004).

85. Gray, R. D.; Chaires, J. B. “Linkage of cation binding and folding in human telomeric quadruplex DNA”. *Biophys. Chem.* 159, 205-209 (2011).

86. Pors Nielsen, S. “The biological role of strontium”. *Bone* 35, 583-588 (2004).

87. Ives, H. E.; Rector, F. C. Jr. “Proton Transport and Cell Function”. *J. Clin. Invest.* 73, 285-290 (1984).

88. Elliot, D.; Ladomery, M. *Molecular Biology of RNA* (Oxford University Press, New York, USA, 2011).

89. Versieck, J.; Barbier, F.; Speecke, A.; Hoste, J. “Manganese, Copper, and Zinc Concentrations in Serum and Packed Blood Cells During Acute Hepatitis, Chronic Hepatitis, and Posthepatitic Cirrhosis”. *Clin. Chem.* 20, 1141-1145 (1974).

90. Ma, J.; Haldar, S.; Khan, M. A.; Sharma, S. D.; Merrick, W. C.; Theil, E. C.; Goss, D. J. *“*Fe^{2+} binds iron responsive element-RNA, selectively changing protein-binding affinities and regulating mRNA repression and activation”. *P. Natl. Acad. Sci. USA* 109, 8417-8422 (2012).

91. Hsiao, C.; Chou, I.-C.; Okafor, C. D.; Bowman, J. C.; O’Neill, E. B.; Athavale, S. S.; Petrov, A. S.; Hud, N. V.; Wartell, R. M.; Harvey, S. C.; Williams, L. D. *“*RNA with iron(II) as a cofactor catalyses electron transfer”. *Nat. Chem.* 5, 525-528 (2013).

92. Kobayashi, M.; Shimizu, S. “Cobalt proteins”. *Eur. J. Biochem.* 261, 1-9 (1999).

93. Abolhasani, J.; Hassanzadeh, J.; Ghorbani-Kalhor, E.; Saeedi, Z. “Fluorescence Quenching of CdS Quantum Dots and Its Application to Determination of Copper and Nickel Contamination in Well and Dam Water”. *J.C.H.R.* 5, 145-154 (2015).

94. Lyons, T. J.; Eide, D. J. “Transport and Storage of Metal Ions in Biology”. In *Biological Inorganic Chemistry: Structure and Reactivity,* edited by I. Bertini, H. B. Gray, E. I. Stiefel & J. S. Valentine (University Science Books, USA, 2007).

95. Maret, W. “Analyzing free zinc(II) ion concentrations in cell biology with fluorescent chelating molecules”. *Metallomics* 7, 202-211 (2015).

96. Russo, A. J.; Devito, R. “Analysis of Copper and Zinc Plasma Concentration and the Efficacy of Zinc Therapy in Individuals with Asperger's Syndrome, Pervasive Developmental Disorder Not Otherwise Specified (PDD-NOS) and Autism”. *Biomarker Insights* 6, 127-133 (2011).

97. Ussing, H. H. *The Alkali Metal Ions in Biology*. (Springer, 1959).

98. Häussinger, D. “Ammonia, urea production and pH regulation”. In *Textbook of Hepatology.* 3rd edition, edited by J. Rodés, J.-P. Benhamou, A. Blei, J. Reichen & M. Rizzetto (Wiley-Blackwell, Oxford, UK, 2007).

99. Essen, L.-O.; Perisic, O.; Lynch, D. E.; Katan, M.; Williams, R. L. “A Ternary Metal Binding Site in the C2 Domain of Phosphoinositide-Specific Phospholipase C-d1”. *Biochemistry* 36, 2753-2762 (1997).

100. Herzfeld, J.; Griffin, R. G.; Haberkorn, R. A. “Phosphorus-31 Chemical-Shift Tensors in Barium Diethyl Phosphate and Urea-Phosphoric Acid: Model Compounds for Phospholipid Head-Group Studies”. *Biochemistry* 17, 2711-2718 (1978).

101. Font, J.; Mackay, J. P. “Beyond DNA: zinc finger domains as RNA-binding modules”. *Methods Mol. Biol.* 649, 479-491 (2010).

102. Williams, R.; Ryves, W. J.; Dalton, E. C.; Eickholt, B.; Shaltiel, G.; Agam, G.; Harwood, A. J. *“*A molecular cell biology of lithium”. *Biochem. Soc. T.* 32, 799-802 (2004).

103. Milne, D. B.; Sims, R. L.; Ralston, N. “Manganese content of the cellular components of blood”. *Clin. Chem.* 36, 450-452 (1990).

104. Angelova, M.; Asenova, S.; Nedkova, V.; Koleva-Kolarova, R. “Copper in the human organism”. *T. J S.* 9, 88-98 (2011).

105. Watson, J. D.; Crick, F. H. C. “Molecular structure of nucleic acids”. *Nature* 171, 73-738 (1953).

106. Zhao, J.; Bacolla, A.; Wang, G.; Vasquez, K. M. “Non-B DNA structure-induced genetic instability and evolution”. *Cell. Mol. Life Sci.* 67, 43-62 (2010).

107. Gellert, M.; Lipsett, M. N.; Davies, D. R. “Helix formation by guanylic acid”. *P. Natl. Acad. Sci. USA* 48, 2013-2018 (1962).

108. Huppert, J. L.; Balasubramanian, S. “Prevalence of quadruplexes in the human genome”. *Nucleic Acids Res.* 33, 2908-2916 (2005).

109. Todd, A. K.; Johnston, M.; Neidle, S. “Highly prevalent putative quadruplex sequence motifs in human DNA”. *Nucleic Acids Res.* 33, 2901-2907 (2005).

110. Jayaraj, G. G.; Pandey, S.; Scaria, V.; Maiti, S. “Potential G-quadruplexes in the human long non-coding transcriptome”. *RNA Biol.* 9, 81-86 (2012).

111. Huppert, J. L.; Balasubramanian, S. “G-quadruplexes in promoters throughout the human genome”. *Nucleic Acids Res.* 35, 406-413 (2007).

112. Kikin, O.; Zappala, Z.; D'Antonio, L.; Bagga, P. S. “GRSDB2 and GRS_UTRdb: databases of quadruplex forming G-rich sequences in pre-mRNAs and mRNAs”. *Nucleic Acids Res.* 36, D141-8 (2008).

113. Biffi, G.; Tannahill, D.; McCafferty, J.; Balasubramanian, S. “Quantitative visualization of DNA G-quadruplex structures in human cells”. *Nat. Chem.* 5, 182-186 (2013).

114. Biffi, G.; Di Antonio, M.; Tannahill, D.; Balasubramanian, S. “Visualization and selective chemical targeting of RNA G-quadruplex structures in the cytoplasm of human cells”. *Nat. Chem.* 6, 75-80 (2014).

115. Bang, I. “Untersuchungen über die Guanylsäure”. *Biochem. Z.* 26, 293-311 (1910).

116. Fresco, J. R.; Massoulié, J. “Helix-coil transition of polyriboguanylic acid”. *J. Am. Chem. Soc.* 85, 1352-1353 (1963).

117. Smargiasso, N.; Rosu, F.; Hsia, W.; Colson, P.; Shammel Baker, E.; Bowers, M. T.; De Pauw, E.; Gabelica, V. *“*G-quadruplex DNA assemblies: loop length, cation identity, and multimer formation”. *J. Am. Chem. Soc.* 130, 10208-10216 (2008).

118. Risitano, A.; Fox, K. R. “Stability of intramolecular DNA quadruplexes: comparison with DNA duplexes”. *Biochemistry* 42, 6507-6513 (2003).

119. Guédin, A.; Gros, J.; Alberti, P.; Mergny, J.-L. “How long is too long? Effects of loop size on G-quadruplex stability”. *Nucleic Acids Res.* 38, 7858-7868 (2010).

120. Zhang, A. Y. Q.; Bugaut, A.; Balasubramanian, S. “A sequence-independent analysis of the loop length dependence of intramolecular RNA G-quadruplex stability and topology”. *Biochemistry* 50, 7251-7258 (2011).

121. Parkinson, G. N.; Lee, M. P. H.; Neidle, S. “Crystal structure of parallel quadruplexes from human telomeric DNA”. *Nature* 417, 876-880 (2002).

122. Campbell, N.; Collie, G. W.; Neidle, S. “Crystallography of DNA and RNA G-Quadruplex Nucleic Acids and Their Ligand Complexes”. *Curr. Protoc. Nucleic Acid Chem.* 50, 17.6.1-17.6.22 (2012).

123. Marathias, V. M.; Bolton, P. H. “Structures of the potassium-saturated, 2:1, and intermediate, 1:1, forms of a quadruplex DNA”. *Nucleic Acids Res.* 28, 1969-1977 (2000).

124. Mao, X.; Marky, L. A.; Gmeiner, W. H. “NMR structure of the thrombin-binding DNA aptamer stabilized by Sr^{2+}”. *J. Biomol. Struct. Dyn.* 22, 25-33 (2004).

125. Wang, Y.; Patel, D. J. “Solution structure of the human telomeric repeat d[$AG_3(T_2AG_3)_3$] G-tetraplex”. *Structure* 1, 263-282 (1993).

126. Halder, K.; Hartig, J. S. “RNA Quadruplexes”. *Met. Ions Life Sci.* 9, 125-139 (2011).

127. Tang, C.-F.; Shafer, R. H. “Engineering the quadruplex fold: nucleoside conformation determines both folding topology and molecularity in guanine quadruplexes”. *J. Am. Chem. Soc.* 128, 5966-5973 (2006).

128. Joachimi, A.; Benz, A.; Hartig, J. S. “A comparison of DNA and RNA quadruplex structures and stabilities”. *Bioorgan. Med. Chem.* 17, 6811-6815 (2009).

129. Collie, G. W.; Parkinson, G. N.; Neidle, S.; Rosu, F.; De Pauw, E.; Gabelica, V. *“*Electrospray mass spectrometry of telomeric RNA (TERRA) reveals the formation of stable multimeric G-quadruplex structures”. *J. Am. Chem. Soc.* 132, 9328-9334 (2010).

130. Lee, M. P H; Parkinson, G. N.; Hazel, P.; Neidle, S. “Observation of the coexistence of sodium and calcium ions in a DNA G-quadruplex ion channel”. *J. Am. Chem. Soc.* 129, 10106-10107 (2007).

131. Lane, A. N.; Chaires, J. B.; Gray, R. D.; Trent, J. O. “Stability and kinetics of G-quadruplex structures”. *Nucleic Acids Res.* 36, 5482-5515 (2008).

132. Lipps, H. J.; Rhodes, D. “G-quadruplex structures: in vivo evidence and function”. *Trends Cell Biol.* 19, 414-422 (2009).

133. Eddy, J.; Maizels, N. “Gene function correlates with potential for G4 DNA formation in the human genome”. *Nucleic Acids Res.* 34, 3887-3896 (2006).

134. Arthanari, H.; Bolton, P. H. “Functional and dysfunctional roles of quadruplex DNA in cells”. *Chem. Biol.* 8, 221-230 (2001).

135. Schaffitzel, C.; Berger, I.; Postberg, J.; Hanes, J.; Lipps, H. J.; Plückthun, A. *“*In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with Stylonychia lemnae macronuclei”. *P. Natl. Acad. Sci. USA* 98, 8572-8577 (2001).

136. Williamson, J. R.; Raghuraman, M. K.; Cech, T. R. “Monovalent cation-induced structure of telomeric DNA: The G-quartet model”. *Cell* 59, 871-880 (1989).

137. Lee, J. Y.; Okumus, B.; Kim, D. S.; Ha, T. “Extreme conformational diversity in human telomeric DNA”. *P. Natl. Acad. Sci. USA* 102, 18938-18943 (2005).

138. Neidle, S.; Parkinson, G. N. “The structure of telomeric DNA”. *Curr. Opin. Struc. Biol.* 13, 275-283 (2003).

139. Ambrus, A.; Chen, D.; Dai, J.; Bialis, T.; Jones, R. A.; Yang, D. *“*Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution”. *Nucleic Acids Res.* 34, 2723-2735 (2006).

140. Li, J.; Correia, J. J.; Wang, L.; Trent, J. O.; Chaires, J. B. “Not so crystal clear: the structure of the human telomere G-quadruplex in solution differs from that present in a crystal”. *Nucleic Acids Res.* 33, 4649-4659 (2005).

141. Paeschke, K.; Simonsson, T.; Postberg, J.; Rhodes, D.; Lipps, H. J. “Telomere end-binding proteins control the formation of G-quadruplex DNA structures in vivo”. *Nat. Struct. Mol. Biol.* 12, 847-854 (2005).

142. Zahler, A. M.; Williamson, J. R.; Cech, T. R.; Prescott, D. M. “Inhibition of telomerase by G-quartet DNA structures”. *Nature* 350, 718-720 (1991).

143. Kim, N. W.; Piatyszek, M. A.; Prowse, K. R.; Harley, C. B.; West, M. D.; Ho, P. L. C.; Coviello, G. M.; Wright, W. E.; Weinrich, S. L.; Shay, J. W. "Specific Association of Human Telomerase Activity with Immortal Cells and Cancer". *Science* 266, 2011-2015 (1994).

144. Neidle, S. "Human telomeric G-quadruplex: the current status of telomeric G-quadruplexes as therapeutic targets in human cancer". *FEBS J.* 277, 1118-1125 (2010).

145. Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H. "Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription". *P. Natl. Acad. Sci. USA* 99, 11593-11598 (2002).

146. Lim, J. K.; Padgett, C. S.; Hoff, D. D.; von, Rice, W. G.; Darjania, L. "Quarfloxin phase I clinical data and scientific findings supporting the selection of carcinoid/neuroendocrine tumors as the phase II indication". *100th AACR Annual Meeting,* 18-22 (2009).

147. Balasubramanian, S.; Hurley, L. H.; Neidle, S. "Targeting G-quadruplexes in gene promoters: a novel anticancer strategy?" *Nat. Rev. Drug Discov.* 10, 261-275 (2011).

148. Bugaut, A.; Balasubramanian, S. "5'-UTR RNA G-quadruplexes: translation regulation and targeting". *Nucleic Acids Res.* 40, 4727-4741 (2012).

149. Huppert, J. L.; Bugaut, A.; Kumari, S.; Balasubramanian, S. "G-quadruplexes: the beginning and end of UTRs". *Nucleic Acids Res.*36, 6260-6268 (2008).

150. Agarwala, P.; Pandey, S.; Maiti, S. "The tale of RNA G-quadruplex". *Org. Biomol. Chem.* 13, 5570-5585 (2015).

151. Kumari, S.; Bugaut, A.; Huppert, J. L.; Balasubramanian, S. "An RNA G-quadruplex in the 5' UTR of the NRAS proto-oncogene modulates translation". *Nat. Chem. Biol.* 3, 218-221 (2007).

152. Bugaut, A.; Rodriguez, R.; Kumari, S.; Hsu, S.-T. D.; Balasubramanian, S. "Small molecule-mediated inhibition of translation by targeting a native RNA G-quadruplex". *Org. Biomol. Chem.* 8, 2771 (2010).

153. Kumari, S.; Bugaut, A.; Balasubramanian, S. "Position and Stability Are Determining Factors for Translation Repression by an RNA G-Quadruplex-Forming Sequence within the 5′ UTR of the NRAS Proto-oncogene". *Biochemistry* 47, 12664-12669 (2008).

154. Arora, A.; Dutkiewicz, M.; Scaria, V.; Hariharan, M.; Maiti, S.; Kurreck, J. "Inhibition of translation in living eukaryotic cells by an RNA G-quadruplex motif". *RNA* 14, 1290-1296 (2008).

155. Weng, H.-Y.; Huang, H.-L.; Zhao, P.-P.; Zhou, H.; Qu, L.-H. "Translational repression of cyclin D3 by a stable G-quadruplex in its 5' UTR: implications for cell cycle regulation". *RNA Biol.* 9, 1099-1109 (2012).

156. Morris, M. J.; Basu, S. "An Unusually Stable G-Quadruplex within the 5′-UTR of the MT3 Matrix Metalloproteinase mRNA Represses Translation in Eukaryotic Cells". *Biochemistry* 48, 5313-5319 (2009).

157. Gomez, D.; Guédin, A.; Mergny, J.-L.; Salles, B.; Riou, J.-F.; Teulade-Fichou, M.-P.; Calsou, P. "A G-quadruplex structure within the 5'-UTR of TRF2 mRNA represses translation in human cells". *Nucleic Acids Res.* 38, 7187-7198 (2010).

158. Balkwill, G. D.; Derecka, K.; Garner, T. P.; Hodgman, C.; Flint, A. P. F.; Searle, M. S. "Repression of translation of human estrogen receptor alpha by G-quadruplex formation". *Biochemistry* 48, 11487-11495 (2009).

159. Morris, M. J.; Negishi, Y.; Pazsint, C.; Schonhoft, J. D.; Basu, S. "An RNA G-quadruplex is essential for cap-independent translation initiation in human VEGF IRES". *J. Am. Chem. Soc.* 132, 17831-17839 (2010).

160. Wolfe, A. L.; Singh, K.; Zhong, Y.; Drewe, P.; Rajasekhar, V. K.; Sanghvi, V. R.; Mavrakis, K. J.; Jiang, M.; Roderick, J. E.; Van der Meulen, J.; Schatz, J. H.; Rodrigo, C. M.; Zhao, C.; Rondou, P.; de Stanchina, E.; Teruya-Feldstein, J.; Kelliher, M. A.; Speleman, F.; Porco, J. A. Jr.; Pelletier, J.; Rätsch, G.; Wendel, H.-G. "RNA G-quadruplexes cause eIF4A-dependent oncogene translation in cancer". *Nature* 513, 65-70 (2014).

161. Agarwala, P.; Pandey, S.; Mapa, K.; Maiti, S. "The G-quadruplex augments translation in the 5' untranslated region of transforming growth factor β2". *Biochemistry* 52, 1528-1538 (2013).

162. Christiansen, J.; Kofod, M.; Nielsen, F. C. "A guanosine quadruplex and two stable hairpins flank a major cleavage site in insulin-like growth factor II mRNA". *Nucleic Acids Res.* 22, 5709-5716 (1994).

163. Sundquist, W. I.; Heaphy, S. "Evidence for interstrand quadruplex formation in the dimerization of human immunodeficiency virus 1 genomic RNA". *P. Natl. Acad. Sci. USA* 90, 3393-3397 (1993).

164. Murat, P.; Zhong, J.; Lekieffre, L.; Cowieson, N. P.; Clancy, J. L.; Preiss, T.; Balasubramanian, S.; Khanna, R.; Tellan, J. "G-quadruplexes regulate Epstein-Barr virus-encoded nuclear antigen 1 mRNA translation". *Nat. Chem. Biol.* 10, 358-364 (2014).

165. Collie, G. W.; Parkinson, G. N. "The application of DNA and RNA G-quadruplexes to therapeutic medicines". *Chem.Soc. Rev.* 40, 5867-5892 (2011).

166. Kaucher, M. S.; Harrell, W. A.; Davis, J. T. "A unimolecular G-quadruplex that functions as a synthetic transmembrane Na^+ transporter". *J. Am. Chem. Soc.* 128, 38-39 (2006).

167. Sen, D.; Gilbert, W. "A sodium-potasium switch in the formation of four-stranded G4-DNA". *Nature* 344, 410-414 (1990).

168. Liu, W.; Fu, Y.; Zheng, B.; Cheng, S.; Li, W.; Lau, T.-C.; Liang, H. "Kinetics and mechanism of conformational changes in a G-quadruplex of thrombin-binding aptamer induced by Pb^{2+}". *J. Phys. Chem. B* 115, 13051-13056 (2011).

169. Moriwaki, H. "Complexes of cadmium ion with guanine bases detected by electrospray ionization mass spectrometry". *J. Mass Spectrom.* 38, 321-327 (2003).

170. a) Monchaud, D.; Teulade-Fichou, M.-P. "A hitchhiker's guide to G-quadruplex ligands". *Org. Biomol. Chem.* 6, 627–636 (2008); b) Georgiades, S. N.; Abd Karim, N. H.; Suntharalingam, K.; Vilar, R. "Interaction of Metal Complexes with G-Quadruplex DNA". *Angew. Chem. Int. Edit.* 49, 4020-4034 (2010).

171. Shivalingam, A.; Izquierdo, M. A.; Le Marois, A.; Vysniauskas, A.; Suhling, K.; Kuimova, M. K.; Vilar, R. "The interactions between a small molecule and G-quadruplexes are visualized by fluorescence lifetime imaging microscopy". *Nat. Commun.* 6, 8178 (2015).

172. Di Antonio, M.; Biffi, G.; Mariani, A.; Raiber, E.-A.; Rodriguez, R.; Balasubramanian, S. "Selective RNA versus DNA G-quadruplex targeting by in situ click chemistry". *Angew. Chem. Int. Edit.* 51, 11073-11078 (2012).

173. Campbell, N. H.; Patel, M.; Tofa, A. B.; Ghosh, R.; Parkinson, G. N.; Neidle, S. "Selectivity in ligand recognition of G-quadruplex loops". *Biochemistry* 48, 1675-1680 (2009).

174. Collie, G. W.; Sparapani, S.; Parkinson, G. N.; Neidle, S. "Structural basis of telomeric RNA quadruplex--acridine ligand recognition". *J. Am. Chem. Soc.* 133, 2721-2728 (2011).

175. Campbell, N. H.; Abd Karim, N. H.; Parkinson, G. N.; Gunaratnam, M.; Petrucci, V.; Todd, A. K.; Vilar, R.; Neidle, S. "Molecular basis of structure-activity relationships between salphen metal complexes and human telomeric DNA quadruplexes". *J. Med. Chem.* 55, 209-222 (2012).

176. Ren, J.; Chaires, J. B. “Sequence and Structural Selectivity of Nucleic Acid Binding Ligands”. *Biochemistry* 38, 16067-16075 (1999).

177. Sabharwal, N. C.; Savikhin, V.; Turek-Herman, J. R; Nicoludis, J. M.; Szalai, V. A.; Yatsunyk, L. A. “N-methylmesoporphyrin IX fluorescence as a reporter of strand orientation in guanine quadruplexes”. *FEBS J.* 281, 1726-1737 (2014).

178. Wu, X.; Maizels, N. “Substrate-specific inhibition of RecQ helicase”. *Nucleic Acids Res.* 29, 1765-1771 (2001).

179. Kim, M.-Y.; Gleason-Guzman, M.; Izbicka, E.; Nishioka, D.; Hurley, L. H. “Biological effects of telomestatin and TMPyP4 can be attributed to their selectivity for interaction with intramolecular or intermolecular G-quadruplex structures”. *Cancer Res.* 63, 3247-3256 (2003).

180. Bejugam, M.; Sewitz, S.; Shirude, P. S.; Rodriguez, R.; Shahid, R.; Balasubramanian, S. “Trisubstituted isoalloxazines as a new class of G-quadruplex binding ligands: small molecule regulation of c-kit oncogene expression”. *J. Am. Chem. Soc.* 129, 12926-12927 (2007).

181. Gonçalves, D. P. N.; Rodriguez, R.; Balasubramanian, S.; Sanders, J. K M. “Tetramethylpyridiniumporphyrazines--a new class of G-quadruplex inducing and stabilising ligands”. *Chem. Commun.* 4685-4687 (2006).

182. Bertrand, H.; Monchaud, D.; De Cian, A.; Guillot, R.; Mergny, J.-L.; Teulade-Fichou, M.-P. “The importance of metal geometry in the recognition of G-quadruplex-DNA by metal-terpyridine complexes”. *Org. Biomol. Chem.* 5, 2555 (2007).

183. Podbevsek, P.; Hud, N. V.; Plavec, J. “NMR evaluation of ammonium ion movement within a unimolecular G-quadruplex in solution”. *Nucleic Acids Res.* 35, 2554-2563 (2007).

184. Collie, G.; Reszka, A. P.; Haider, S. M.; Gabelica, V.; Parkinson, G. N.; Neidle, S. “Selectivity in small molecule binding to human telomeric RNA and DNA quadruplexes”. *Chem. Commun.* 7482-7484 (2009).

185. Cousins, A. R. O.; Ritson, D.; Sharma, P.; Stevens, M. F. G.; Moses, J. E.; Searle, M. S. “Ligand selectivity in stabilising tandem parallel folded G-quadruplex motifs in human telomeric DNA sequences”. *Chem. Commun.* 50, 15202-15205 (2014).

186. Rzuczek. S. G.; Pilch, D. S.; Liu, A.; La Voie, E. J.; Rice, J. E. “Macrocyclic pyridyl polyoxazoles: selective RNA and DNA G-quadruplex ligands as antitumor agents”. *J. Med. Chem.* 53, 3632-3644 (2010).

187. Garner, T. P.; Williams, H. E. L.; Gluszyk, K. I.; Roe, S.; Oldham, N. J.; Stevens, M. F. G.; Moses, J. E.; Searle, M. S. “Selectivity of small molecule ligands for parallel and anti-parallel DNA G-quadruplex structures”. *Org. Biomol. Chem.* 7, 4194-4200 (2009).

188. Cheong, C.; Moore, P. B. “Solution Structure of an Unusually Stable RNA Tetraplex Containing G- and U-Quartet Structures”. *Biochemistry* 31, 8406-8414 (1992).

189. a) Parikh, C.; Subrahmanyam, R.; Ren, R. “Oncogenic NRAS rapidly and efficiently induces CMML- and AML-like diseases in mice”. *Blood* 108, 2349-2357 (2006); b) Lázár, V.; Ecsedi, S.; Szöllösi, A. G.; Tóth, R.; Vízkeleti, L.; Rákosy, Z.; Bégány, A.; Ádány, R.; Balász, M. “Characterization of candidate gene copy number alterations in the 11913 region along with *BRAF* and *NRAS* mutations in human melanoma”. *Mod. Pathol.* 22, 1367-1378 (2009).

190. Xu, Y.; Suzuki, Y.; Ito, K.; Komiyama, M. “Telomeric repeat-containing RNA structure in living cells”. *P. Natl. Acad. Sci. USA* 107, 14579-14584 (2010).

191. Mergny, J.-L.; Phan, A. T.; Lacroix, L. “Following G-quartet formation by UV-spectroscopy”. *FEBS Lett.* 435, 74-78 (1998).

192. Mergny, J.-L.; Lacroix, L. “Analysis of Thermal Melting Curves”. *Oligonucleotides* 13, 515-537 (2003).

193. Woody, R. W. “Circular dichroism”. In *Biochemical Spectroscopy* (Elsevier 1995).

194. Mergny, J.-L.; Lacroix, L. “UV Melting of G-Quadruplexes”. *Curr. Protoc. Nucleic Acid Chem.* Chapter 17, Unit 17.1 (2009).

195. Małgowska, M.; Gudanis, D.; Teubert, A.; Dominiak, G.; Gdaniec, Z. “How to study G-quadruplex structures”. *BioTechnologia* 93, 381-390 (2012).

196. Mergny, J.-L.; Li, J.; Lacroix, L.; Amrane, S.; Chaires, J. B. “Thermal difference spectra: a specific signature for nucleic acid structures”. *Nucleic Acids Res.* 33, e138 (2005).

197. Rio, D. C.; Ares, M.; Hannon, G. J.; Nilsen, T. W. “Polyacrylamide gel electrophoresis of RNA”. *Cold Spring Harb. Protoc.* (2010).

198. Adrian, M.; Ang, D. J.; Lech, C. J.; Heddi, B.; Nicolas, A.; Phan, A. T. “Structure and conformational dynamics of a stacked dimeric G-quadruplex formed by the human CEB1 minisatellite”. *J. Am. Chem. Soc.* 136, 6297-6305 (2014).

199. Goodlett, D. R.; Camp II, D. G.; Hardin, C. C.; Corregan, M.; Smith, R. D. “Direct observation of a DNA quadruplex by electrospray ionization mass spectrometry”. *Biol. Mass Spectrom.* 22, 181-183 (1993).

200. Rosu, F.; Gabelica, V.; Houssier, C.; Colson, P.; de Pauw, E. “Triplex and quadruplex DNA structures studied by electrospray mass spectrometry”. *Rapid Commun. Mass Sp.* 16, 1729-1736 (2002).

201. Rosu, F.; Gabelica, V.; Poncelet, H.; de Pauw, E. “Tetramolecular G-quadruplex formation pathways studied by electrospray mass spectrometry”. *Nucleic Acids Res.*38, 5217-5225 (2010).

202. Rosu, F.; de Pauw, E.; Gabelica, V. “Electrospray mass spectrometry to study drug-nucleic acids interactions”. *Biochimie* 90, 1074-1087 (2008).

203. Yuan, G.; Zhang, Q.; Zhou, J.; Li, H. “Mass spectrometry of G-quadruplex DNA: formation, recognition, property, conversion, and conformation”. *Mass Spectrom. Rev.* 30, 1121-1142 (2011).

204. Goldburg, W. I. “Dynamic light scattering”. *Am. J. Phys.* 67, 1152 (1999).

205. Johannsen, S.; Paulus, S.; Düpre, N.; Müller, J.; Sigel, R. K O. “Using in vitro transcription to construct scaffolds for one-dimensional arrays of mercuric ions”. *J. Inorg. Biochem.* 102, 1141-1151 (2008).

206. Adrian, M.; Heddi, B.; Phan, A. T. “NMR spectroscopy of G-quadruplexes”. *Methods* 57, 11-24 (2012).

207. Cohen, Y.; Avram, L.; Frish, L. “Diffusion NMR spectroscopy in supramolecular and combinatorial chemistry: an old parameter--new insights”. *Angew. Chem. Int. Edit.* 44, 520-554 (2005).

208. Ambrus, A.; Yang, D. “Diffusion-ordered nuclear magnetic resonance spectroscopy for analysis of DNA secondary structural elements”. *Anal. Biochem.* 367, 56-67 (2007).

209. Sket, P.; Plavec, J. “Not all G-quadruplexes exhibit ion-channel-like properties: NMR study of ammonium ion (non)movement within the $d(G_3T_4G_4)_2$ quadruplex”. *J. Am. Chem. Soc.* 129, 8794-8800 (2007).

210. Šket, P.; Plavec, J. “Tetramolecular DNA Quadruplexes in Solution: Insights into Structural Diversity and Cation Movement”. *J. Am. Chem. Soc.* 132, 12724-12732 (2010).

211. Wu, G.; Wong, A.; Gan, Z.; Davis, J. T. “Direct detection of potassium cations bound to G-quadruplex structures by solid-state 39K NMR at 19.6 T”. *J. Am. Chem. Soc.* 125, 7182-7183 (2003).

212. Wong, A.; Ida, R.; Wu, G. “Direct NMR detection of the "invisible" alkali metal cations tightly bound to G-quadruplex structures”. *Biochem. Bioph. Res. Co.* 337, 363-366 (2005).

213. Ida, R.; Wu, G. “Direct NMR Detection of Alkali Metal Ions Bound to G-Quadruplex DNA”. *J. Am. Chem. Soc.* 130, 3590-602 (2008).

214. Miyoshi, D.; Nakao, A.; Sugimoto, N. “Molecular Crowding Regulates the Structural Switch of the DNA G-Quadruplex”. *Biochemistry* 41, 15017-15024 (2002).

215. Miyoshi, D.; Karimata, H.; Sugimoto, N. “Hydration regulates thermodynamics of G-quadruplex formation under molecular crowding conditions”. *J. Am. Chem. Soc.* 128, 7957-7963 (2006).

216. Fedoroff, O. Y.; Salazar, M.; Han, H.; Chemeris, V. V.; Kerwin, S. M.; Hurley, L. H. “NMR-Based Model of a Telomerase-Inhibiting Compound Bound to G-Quadruplex DNA”. *Biochemistry* 37, 12367-12374 (1998).

217. Dai, J.; Chen, D.; Jones, R. A.; Hurley, L. H.; Yang, D. “NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region”. *Nucleic Acids Res.* 34, 5133-5144 (2006).

218. Ha, T.; Enderle, Th.; Ogletree, D. F.; Chemla, D. S.; Selvin, P. R.; Weiss, S. “Probing the interaction between two single molecules: Fluorescence resonance energy transfer between a single donor and a single acceptor”. *P. Natl. Acad. Sci. USA* 93, 6264-6268 (1996).

219. Roy, R.; Hohng, S.; Ha, T. “A practical guide to single-molecule FRET”. *Nat. Methods* 5, 507-516 (2008).

220. Joo, C.; Balci, H.; Ishitsuka, Y.; Buranachai, C.; Ha, T. “Advances in single-molecule fluorescence methods for molecular biology” *Annu. Rev. Biochem.* 77, 51-76 (2008).

221. Shirude, P. S.; Balasubramanian, S. “Single molecule conformational analysis of DNA G-quadruplexes”. *Biochimie* 90, 1197-1206 (2008).

222. Steiner, M.; Karunatilaka, K. S.; Sigel, R. K. O.; Rueda, D. “Single-molecule studies of group II intron ribozymes”. *P. Natl. Acad. Sci. USA* 105, 13853-13858 (2008).

223. Förster, T. “Zwischenmolekulare Energiewanderung und Fluoreszenz”. *Ann. Phys.* 437, 55-75 (1948).

224. Joo, C.; Ha, T. “Single-molecule FRET with total internal reflection microscopy”. *Cold Spring Harb. Protoc.* (2012).

225. Schmitz, A. G.; Zelger-Paulus, S.; Gasser, G.; Sigel, R. K O. “Strategy for Internal Labeling of Large RNAs with Minimal Perturbation by Using Fluorescent PNA”. *Chembiochem* 16, 1302-1306 (2015).

226. Benesch, R. E.; Benesch, R. “Enzymatic Removal of Oxygen for Polarography and Related Methods”. *Science* 118, 447-448 (1953).

227. Joo, C.; Ha, T. “Preparing sample chambers for single-molecule FRET”. *Cold Spring Harb. Protoc.* 2012, 1104-1108 (2012).

228. Zhao, R.; Rueda, D. “RNA folding dynamics by single-molecule fluorescence resonance energy transfer”. *Methods* 49, 112-117 (2009).

229. Sunney Xie, X.; Yu, J.; Yang, W. Y. “Living Cells as Test Tubes”. *Science* 312, 228-230 (2006).

230. Sako, Y. “Imaging single molecules in living cells for systems biology”. *Mol. Syst. Biol.* 2, 56 (2006).

231. Włodarczyk, A.; Grzybowski, P.; Patkowski, A.; Dobek, A. “Effect of ions on the polymorphism, effective charge, and stability of human telomeric DNA”. Photon correlation spectroscopy and circular dichroism studies. *J. Phys. Chem. B* 109, 3594-3605 (2005).

232. Miyoshi, D.; Nakao, A.; Toda, T.; Sugimoto, N. “Effect of divalent cations on antiparallel G-quartet structure of d($G_4T_4G_4$)”. *FEBS Lett.* 496, 128-133 (2001).

233. Venczel, E. A.; Sen, D. “Parallel and antiparallel G-DNA structures from a complex telomeric sequence”. *Biochemistry* 32, 6220-6228 (1993).

234. Chen, F. M. “Strontium(2+) facilitates intermolecular G-quadruplex formation of telomeric sequences”. *Biochemistry* 31, 3769-3776 (1992).

235. Hardin, C. C.; Watson, T.; Corregan, M.; Bailey, C. “Cation-dependent transition between the quadruplex and Watson-Crick hairpin forms of d($CGCG_3GCG$)”. *Biochemistry* 31, 833-841 (1992).

236. Kypr, J.; Kejnovská, I.; Renciuk, D.; Vorlícková, M. “Circular dichroism and conformational polymorphism of DNA”. *Nucleic Acids Res.* 37, 1713-1725 (2009).

237. Hong, J.; Capp, M. W.; Anderson, C. F.; Saecker, R. M.; Felitsky, D. J.; Anderson, M. W.; Record, M. T. Jr. “Preferential interactions of glycine betaine and of urea with DNA: implications for DNA hydration and for effects of these solutes on DNA stability”.

238. Gray, R. D.; Chaires, J. B. “Isothermal folding of G-quadruplexes”. *Methods* 57, 47-55 (2012).

239. Kandegedara, A.; Rorabacher, D. B. “Noncomplexing Tertiary Amines as "Better" Buffers Covering the Range of pH 3-11. Temperature Dependence of Their Acid Dissociation Constants”. *Anal. Chem.* 71, 3140-3144 (1999).

240. Fukada, H.; Takahashi, K. “Enthalpy and heat capacity changes for the proton dissociation of various buffer components in 0.1 M potassium chloride”. *Proteins* 33, 159-166 (1998).

241. Phan, A. T.; Mergny, J.-L. “Human telomeric DNA: G-quadruplex, i-motif and Watson-Crick double helix”. *Nucleic Acids Res.* 30, 4618-4625 (2002).

242. González-Rodríguez, D.; van Dongen, J. L. J.; Lutz, M.; Spek, A. L.; Schenning, A. P. H. J.; Meijer, E. W. *“*G-quadruplex self-assembly regulated by Coulombic interactions”. *Nat. Chem.* 1, 151-155 (2009).

243. Vorlíčková, M.; Kejnovska, I.; Sagi, J.; Renciuk, D.; Bednarova, K.; Motlova, J.; Kypr, J. “Circular dichroism and guanine quadruplexes”. *Methods* 57, 64-75 (2012).

244. Rieger, L.; Langergraber, G.; Thomann, M.; Fleischmann, N.; Siegrist, H. “Spectral in-situ analysis of NO_2, NO_3, COD, DOC and TSS in the effluent of a WWTP”. *Water Sci. Technol.* 50, 143-152 (2004).

245. Trajkovski, M.; Sket, P.; Plavec, J. “Cation localization and movement within DNA thrombin binding aptamer in solution”. *Org. Biomol. Chem.* 7, 4677-4684 (2009).

246. Ji, X.; Sun, H.; Zhou, H.; Xiang, J.; Tang, Y.; Zhao, C. “Research progress of RNA quadruplex”. *Nucleic Acid Ther.* 21, 185-200 (2011).

247. Kim, B. G.; Shek, Y. L.; Chalikian, T. V. “Polyelectrolyte effects in G-quadruplexes”. *Biophys. Chem.* 184, 95-100 (2013).

248. Davis, J. T. “G-quartets 40 years later: from 5'-GMP to molecular biology and supramolecular chemistry”. *Angew. Chem. Int. Edit.* 43, 668-698 (2004).

249. Zhang, D.; Huang, T.; Lukeman, P. S.; Paukstelis, P. J. “Crystal structure of a DNA/Ba2+ G-quadruplex containing a water-mediated C-tetrad”. *Nucleic Acids Res.* 42, 13422-13429 (2014).

250. Pan, B.; Xiong, Y.; Shi, K.; Deng, J.; Sundaralingam, M. “Crystal Structure of an RNA Purine-Rich Tetraplex Containing Adenine Tetrads”. *Structure* 11, 815-823 (2003).

251. Campbell, N. H.; Neidle, S. “G-quadruplexes and metal ions”. *Met. Ions Life Sci.* 10, 119-134 (2012).

252. Deng, J.; Xiong, Y.; Sundaralingam, M. “X-ray analysis of an RNA tetraplex (UGGGGU)4 with divalent Sr^{2+} ions at subatomic resolution (0.61 A)”. *P. Natl. Acad. Sci. USA* 98, 13665-13670 (2001).

253. Ganai, S. A.; Chishti, H.-T.-N.; Ahmad, J.; Ahmad, S. “Synthesis and Thermal Studies of Polyaniline Stannic Silicate and Its Role in the Removal of Toxic Metal Ions”. *A. J. A. C.* 03, 272-276 (2012).

254. Hanzlik, R. *Inorganic Aspects of Biological and Organic Chemistry* (Academic Press, Inc. 1976).

255. Hanafi, A. “Adsorption of cesium, thallium, strontium and cobalt radionuclides using activated carbon”. *J. A. M. S.* 1, 292-300 (2010).

256. Jorgensen, T.; Weatherley, L. “Ammonia removal from wastewater by ion exchange in the presence of organic contaminants”. *Water Res.* 37, 1723-1728 (2003).

257. Richards, L. A.; Richards, B. S.; Schäfer, A. I. “Renewable energy powered membrane technology: Salt and inorganic contaminant removal by nanofiltration/reverse osmosis”. *J. Membrane Sci.* 369, 188-195 (2011).

258. Smith, D. W. “Ionic hydration enthalpies”. *J. Chem. Educ.* 54, 540-542 (1977).

259. Seo, J.; Hong, E. S.; Yoon, H.-J.; Shin, S. K. “Specific and nonspecific bindings of alkaline-earth metal ions to guanine-quadruplex thrombin-binding aptamer DNA”. *Int. J. Mass Spectrom.* 330-332, 262-270 (2012).

260. Wulfsberg, G. *Inorganic Chemistry* (University Science Books, 2000).

261. Boda, A.; Ali, S. M. “From microhydration to bulk hydration of Rb+ metal ion: DFT, MP2 and AIMD simulation study”. *J. Mol. Liq.* 179, 34-45 (2013).

262. Chaudhari, M. I.; Soniat, M.; Rempe, S. B. “Octa-coordination and the Hydrated Ba^{2+} (aq) Ion”. *arXiv.org, Physics, Cornell University Library,* 1-7 (2014).

263. Chang, T.-M.; Dang, L. X. “On rotational dynamics of an NH_4^+ ion in water”. *J. Chem. Phys.* 118, 8813-8820 (2003).

264. Wei, D.; Parkinson, G. N.; Reszka, A. P.; Neidle, S. “Crystal structure of a c-kit promoter quadruplex reveals the structural role of metal ions and water molecules in maintaining loop conformation”. *Nucleic Acids Res.* 40, 4691-4700 (2012).

265. Chowdhury, S.; Bansal, M. “G-Quadruplex Structure Can Be Stable with Only Some Coordination Sites Being Occupied by Cations: A Six-Nanosecond Molecular Dynamics Study”. *J. Phys. Chem. B* 105, 7572-7578 (2001).

266. Lee, J. Y.; Yoon, J.; Kihm, H. W.; Kim, D. S. “Structural diversity and extreme stability of unimolecular Oxytricha nova telomeric G-quadruplex”. *Biochemistry* 47, 3389-3396 (2008).

267. He, F.; Tang, Y.; Wang, S.; Li, Y.; Zhu, D. “Fluorescent amplifying recognition for DNA G-quadruplex folding with a cationic conjugated polymer: a platform for homogeneous potassium detection”. *J. Am. Chem. Soc.* 127, 12343-12346 (2005).

268. Arora, A.; Maiti, S. “Differential biophysical behavior of human telomeric RNA and DNA quadruplex”. *J. Phys. Chem. B* 113, 10515-10520 (2009).

269. Shim, J. W.; Tan, Q.; Gu, L.-Q. “Single-molecule detection of folding and unfolding of the G-quadruplex aptamer in a nanopore nanocavity”. *Nucleic Acids Res.* 37, 972-982 (2009).

270. Qu, K.; Zhao, C.; Ren, J.; Qu, X. “Human telomeric G-quadruplex formation and highly selective fluorescence detection of toxic strontium ions”. *Mol. Biosyst.* 8, 779-782 (2012).

271. Blume, S. W.; Guarcello, V.; Zacharias, W.; Miller, D. M. “Divalent transition metal cations counteract potassium-induced quadruplex assembly of oligo(dG) sequences”. *Nucleic Acids Res.* 25, 617-625 (1997).

272. Monchaud, D.; Yang, P.; Lacroix, L.; Teulade-Fichou, M.-P.; Mergny, J.-L. “A Metal-Mediated Conformational Switch Controls G-Quadruplex Binding Affinity”. *Angew. Chem. Int. Edit.* 120, 4936-4939 (2008).

273. Engelhard, D. M.; Pievo, R.; Clever, G. H. “Reversible stabilization of transition-metal-binding DNA G-quadruplexes”. *Angew. Chem. Int. Edit.* 52, 12843-12847 (2013).

274. Balaratnam, S.; Basu, S. “Divalent cation-aided identification of physico-chemical properties of metal ions that stabilize RNA G-quadruplexes”. *Biopolymers* 103, 376-386 (2015).

275. Wang, Y.-X.; Huang, S.; Draper, D. E. “Structure of a U·U pair within a conserved ribosomal RNA hairpin”. *Nucleic Acids Res.* 24, 2666-2672 (1996).

276. Nakano, S.-i.; Miyoshi, D.; Sugimoto, N. “Effects of molecular crowding on the structures, interactions, and functions of nucleic acids”. *Chem. Rev.*114, 2733-2758 (2014).

277. Heddi, B.; Phan, A. T. “Structure of human telomeric DNA in crowded solution”. *J. Am. Chem. Soc.* 133, 9824-9833 (2011).

278. Buscaglia, R.; Clarke Miller, M.; Dean, W. L.; Gray, R. D.; Lane, A. N.; Trent, J. O.; Chaires, J. B. “Polyethylene glycol binding alters human telomere G-quadruplex structure by conformational selection”. *Nucleic Acids Res.* 41, 7934-7946 (2013).

279. Morris, M. J.; Wingate, K. L.; Silwal, J.; Leeper, T. C.; Basu, S. “The porphyrin TmPyP4 unfolds the extremely stable G-quadruplex in MT3-MMP mRNA and alleviates its repressive effect to enhance translation in eukaryotic cells”. *Nucleic Acids Res.* 40, 4137-4145 (2012).

280. Pandey, S.; Agarwala, P.; Maiti, S. “Effect of loops and G-quartets on the stability of RNA G-quadruplexes”. *J. Phys. Chem. B* 117, 6896-6905 (2013).

281. Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. “Quadruplex DNA: sequence, topology and structure”. *Nucleic Acids Res.* 34, 5402-5415 (2006).

282. Collie, G. W.; Haider, S. M.; Neidle, S.; Parkinson, G. N. “A crystallographic and modelling study of a human telomeric RNA (TERRA) quadruplex”. *Nucleic Acids Res.* 38, 5569-5580 (2010).

283. Martadinata, H.; Phan, A. T. “Structure of propeller-type parallel-stranded RNA G-quadruplexes, formed by human telomeric RNA sequences in K^+ solution”. *J. Am. Chem. Soc.* 131, 2570-2578 (2009).

284. Wong, A.; Wu, G. “Selective binding of monovalent cations to the stacking G-quartet structure formed by guanosine 5'-monophosphate: a solid-state NMR study”. *J. Am. Chem. Soc.* 125, 13895-13905 (2003).

285. Martadinata, H.; Phan, A. T. “Structure of Human Telomeric RNA (TERRA): Stacking of Two G-Quadruplex Blocks in K^+ Solution”. *Biochemistry* 52, 2176-2183 (2013).

286. Mukundan, V. T.; Do, N. Q.; Phan, A. T. “HIV-1 integrase inhibitor T30177 forms a stacked dimeric G-quadruplex structure containing bulges”. *Nucleic Acids Res.* 39, 8984-8991 (2011).

287. Guiset Miserachs, H.; Donghi, D.; Börner, R.; Johannsen, S.; Sigel, R. K. O. “Distinct differences in metal ion specificity of RNA and DNA G-quadruplexes”. *J. Biol. Inorg. Chem.* (2016), *accepted*.

288. Zhang, D.-H.; Fujimoto, T.; Saxena, S.; Yu, H.-Q.; Miyoshi, D.; Sugimoto, N. “Monomorphic RNA G-quadruplex and polymorphic DNA G-quadruplex structures responding to cellular environmental factors”. *Biochemistry* 49, 4554-4563 (2010).

289. Peacock, A. C.; Dingman, C. W. “Resolution of Multiple Ribonucleic Acid Species by Polyacrylamide Gel Electrophoresis”. *Biochemistry* 6, 1818-1827 (1967).

290. Le, H. T.; Clarke Miller, M.; Buscaglia, R.; Dean, W. L.; Holt, P. A.; Chaires, J. B.; Trent, J. O. “Not all G-quadruplexes are created equally: an investigation of the structural polymorphism of the c-Myc G-quadruplex-forming sequence and its interaction with the porphyrin TMPyP4”. *Org. Biomol. Chem.* 10, 9393-9404 (2012).

291. Lubitz, I.; Zikich, D.; Kotlyar, A. “Specific high-affinity binding of thiazole orange to triplex and G-quadruplex DNA”. *Biochemistry* 49, 3567-3574 (2010).

292. Tuma, R. S.; Beaudet, M. P.; Jin, X.; Jones, L. J.; Cheung, C.-Y.; Yue, S.; Singer, V. L. “Characterization of SYBR Gold Nucleic Acid Gel Stain: A Dye Optimized for Use with 300-nm Ultraviolet Transilluminators”. *Anal. Biochem.* 268, 278-288 (1999).

293. Milligan, J. F.; Uhlenbeck, O. C. “Synthesis of small RNAs using T7 RNA polymerase”. In *RNA Processing Part A: General Methods* (Elsevier 1989).

294. Usman, N.; Cedergren, R. “Exploiting the chemical synthesis of RNA”. *Trends Biochem. Sci.* 17, 334-339 (1992).

295. Phan, A. T.; Kuryavyi, V.; Ma, J.-B.; Faure, A.; Andréola, M.-L.; Patel, D. J. “An interlocked dimeric parallel-stranded DNA quadruplex: A potent inhibitor of HIV-1 integrase”. *P. Natl. Acad. Sci. USA* 102, 634-639 (2005).

296. Kejnovská, I.; Kypr, J.; Vorlícková, M. “Oligo(dT) is not a correct native PAGE marker for single-stranded DNA”. *Biochem. Bioph. Res. Co.* 353, 776-779 (2007).

297. Rueda, M.; Luque, F. J.; Orozco, M. “G-quadruplexes can maintain their structure in the gas phase”. *J. Am. Chem. Soc.* 128, 3608-3619 (2006).

298. Deguchi, K.; Masako, I.; Yokokura, T. ; Ogata, I. ; Ito, S.; Mimura, T.; Ostrander, C. “Enhanced mass detection of oligonucleotides using reverse-phase high-performance liquid chromatography/electrospray ionization ion-trap mass spectrometry”. *Rapid Commun. Mass Sp.* 16, 2133-2141 (2002).

299. Phan, A. T.; Do, N. Q. “Engineering of interlocked DNA G-quadruplexes as a robust scaffold”. *Nucleic Acids Res.* 41, 2683-2688 (2013).

300. Gray, R. D.; Buscaglia, R.; Chaires, J. B. “Populated intermediates in the thermal unfolding of the human telomeric quadruplex”. *J. Am. Chem. Soc.* 134, 16834-16844 (2012).

301. Okamoto, K.; Sannohe, Y.; Mashimo, T.; Sugiyama, H.; Terazima, M. “G-quadruplex structures of human telomere DNA examined by single molecule FRET and BrG-substitution”. *Bioorg. Med. Chem.* 16, 6873-6879 (2008).

302. Ying, L.; Green, J. J.; Li, H.; Klenerman, D.; Balasubramanian, S. “Studies on the structure and dynamics of the telomeric G quadruplex by single-molecule fluorescence resonance energy transfer”. *P. Natl. Acad. Sci. USA* 100, 14629-14634 (2003).

303. Shirude, P. S.; Ying, L.; Balasubramanian, S. “Single molecule conformational analysis of the biologically relevant DNA G-quadruplex in the promoter of the proto-oncogene c-MYC”. *Chem. Commun.* 2007-2009 (2008).

304. Shirude, P. S.; Okumus, B.; Ying, L.; Ha, T.; Balasubramanian, S. “Single-molecule conformational analysis of G-quadruplex formation in the promoter DNA duplex of the proto-oncogene c-kit”. *J. Am. Chem. Soc.* 129, 7484-7485 (2007).

305. Okumus, B.; Ha, T. “Real-time observation of G-quadruplex dynamics using single-molecule FRET microscopy”. *Methods Mol. Biol.* 608, 81-96 (2010).

306. Fegan, A.; Shirude, P. S.; Ying, L.; Balasubramanian, S. “Ensemble and single molecule FRET analysis of the structure and unfolding kinetics of the c-kit promoter quadruplexes”. *Chem.Commun.* 46, 946-948 (2010).

307. Tippana, R.; Xiao, W.; Myong, S. “G-quadruplex conformation and dynamics are determined by loop length and sequence”. *Nucleic Acids Res.* 42, 8106-8114 (2014).

308. Long, X.; Stone, M. D. “Kinetic partitioning modulates human telomere DNA G-quadruplex structural polymorphism”. *PLoS ONE* 8, e83420 (2013).

309. Hwang, H.; Buncher, N.; Opresko, P. L.; Myong, S. “POT1-TPP1 regulates telomeric overhang structural dynamics”. *Structure* 20, 1872-1880 (2012).

310. Tanaka, A.; Choi, J.; Kim, S. K.; Majima, T. “Interaction of G-quadruplex with RecA protein studied in bulk phase and at the single-molecule level”. *J. Phys. Chem. B* 117, 6711-6717 (2013).

311. Zhou, R.; Zhang, J.; Bochman, M. L.; Zakian, V. A.; Ha, T. “Periodic DNA patrolling underlies diverse functions of Pif1 on R-loops and G-rich DNA”. *eLIFE* 3, e02190 (2014).

312. Huber, M. D.; Lee, D. C.; Maizels, N. “G4 DNA unwinding by BLM and Sgs1p: substrate specificity and substrate-specific inhibition”. *Nucleic Acids Res.* 30, 3954-3961 (2002).

313. Budhathoki, J. B.; Stafford, E. J.; Yodh, J. G.; Balci, H. “ATP-dependent G-quadruplex unfolding by Bloom helicase exhibits low processivity”. *Nucleic Acids Res.* 43, 5961-5970 (2015).

314. Budhathoki, J. B.; Ray, S.; Urban, V.; Janscak, P.; Yodh, J. G.; Balci, H. “RecQ-core of BLM unfolds telomeric G-quadruplex in the absence of ATP”. *Nucleic Acids Res.* 42, 11528-11545 (2014).

315. Brooks, T. A.; Hurley, L. H. “Targeting MYC Expression through G-Quadruplexes”. *Genes Cancer* 1, 641-649 (2010).

316. Wang, F.; Podell, E. R.; Zaug, A. J.; Yang, Y.; Baciu, P.; Cech, T. R.; Lei, M. “The POT1-TPP1 telomere complex is a telomerase processivity factor”. *Nature* 445, 506-510 (2007).

317. Jankowsky, E. “RNA helicases at work: binding and rearranging”. *Trends Biochem. Sci.* 36, 19-29 (2011).

318. Sissi, C.; Gatto, B.; Palumbo, M. “The evolving world of protein-G-quadruplex recognition: a medicinal chemist's perspective”. *Biochimie* 93, 1219-1230 (2011).

319. Nguyen, G. H.; Tang, W.; Robles, A. I.; Beyer, R. P.; Gray, L. T.; Welsh, J. A.; Schetter, A. J.; Kumamoto, K.; Wang, X. W.; Hickson, I. D.; Maizels, N.; Monnat, R. J. Jr.; Harris, C. C. “Regulation of gene expression by the BLM helicase correlates with the presence of G-quadruplex DNA motifs”. *P. Natl. Acad. Sci. USA* 111, 9905-9910 (2014).

320. Smestad, J.; Maher, L. J. “Putative G-quadruplex forming sequence signatures in genes differentially transcribed upon loss of BLM or WRN helicases”. *bioRxiv* (2015).

321. Wu, W.-Q.; Hou, X.-M.; Li, M.; Dou, S.-X.; Xi, X.-G. “BLM unfolds G-quadruplexes in different structural environments through different mechanisms”. *Nucleic Acids Res.* 43, 4614-4626 (2015).

322. Chatterjee, S.; Zagelbaum, J.; Savitsky, P.; Sturzenegger, A.; Huttner, D.; Janscak, P.; Hickson, I. D.; Gileadi, O.; Rothenberg, E. “Mechanistic insight into the interaction of BLM helicase with intra-strand G-quadruplex structures”. *Nat. Commun.* 5, 5556 (2014).

323. Roy, R.; Hohng, S.; Ha, T. “A practical guide to single-molecule FRET”. *Nat. Methods* 5, 507-516 (2008).

324. Joo, C.; Ha, T. “Labeling DNA (or RNA) for single-molecule FRET”. *Cold Spring Harb. Protoc.* 1005-1008 (2012).

325. Wood, S.; Rueda, D. “Fluorescence Labeling of Nucleic Acids”. In *Encyclopedia of Biophysics,* edited by G. C. Roberts (Springer-Verlag, Berlin Heidelberg, 2013).

326. Lumiprobe. Protocol: NHS Ester Labeling of Amino-Biomolecules. Available at http://www.lumiprobe.com/protocols/nhs-ester-labeling.

327. Solomatin, S.; Herschlag, D. “Methods of site-specific labeling of RNA with fluorescent dyes”. In *Methods in Enzymology.* Part B, edited by D. Herschlag (Elsevier, Amsterdam, Netherlands, 2009).

328. Joo, C.; Ha, T. *Single-Molecules Techniques. A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2008).

329. Walter, N. G. “Probing RNA Structural Dynamics and Function by Fluorescence Resonance Energy Transfer (FRET)”. *Curr. Protoc. Nucleic Acid Chem.* 11, 11.10.1-11.10.20 (2002).

330. Rachwal, P. A.; Fox, K. R. “Quadruplex melting”. *Methods* 43, 291-301 (2007).

331. Qin, H.; Ren, J.; Wang, J.; Luedtke, N. W.; Wang, E. “G-quadruplex-modulated fluorescence detection of potassium in the presence of a 3500-fold excess of sodium ions”. *Anal. Chem.* 82, 8356-8360 (2010).

332. Kapanidis, A. N.; Laurence, T. A.; Lee, N. L.; Margeat, E.; Kong, X.; Weiss, S. “Alternating-laser excitation of single molecules”. *Accounts Chem. Res.* 38**,** 523-533 (2005).

333. Kim, H.-K.; Rasnik, I.; Liu, J.; Ha, T.; Lu, Y. “Dissecting metal ion-dependent folding and catalysis of a single DNAzyme”. *Nat. Chem. Biol.* 3, 763-768 (2007).

334. König, S. L. B.; Hadzic, M.; Fiorini, E.; Börner, R.; Kowerko, D.; Blanckenhorn, W. U.; Sigel, R. K. O. “BOBA FRET: bootstrap-based analysis of single-molecule FRET data”. *PLoS ONE* 8, e84157 (2013).

335. Seo, M.-H.; Lee, T.-S.; Kim, E.; Cho, Y. L.; Park, H.-S.; Yoon, T.-Y.; Kim, H.-S. “Efficient single-molecule fluorescence resonance energy transfer analysis by site-specific dual-labeling of protein using an unnatural amino acid”. *Anal. Chem.* 83, 8849-8854 (2011).

336. Creacy, S. D.; Routh, E. D.; Iwamoto, F.; Nagamine, Y.; Akman, S. A.; Vaughn, J. P. “G4 resolvase 1 binds both DNA and RNA tetramolecular quadruplex with high affinity and is the major source of tetramolecular quadruplex G4-DNA and G4-RNA resolving activity in HeLa cell lysates”. *J. Biol. Chem.* 283, 34626-34634 (2008).

337. Huang, W.; Smaldino, P. J.; Zhang, Q.; Miller, L. D.; Cao, P.; Stadelman, K.; Wan, M.; Giri, B.; Lei, M.; Nagamine, Y.; Vaugh, J. P.; Akman, S. A.; Sui, G. “Yin Yang 1 contains G-quadruplex structures in its promoter and 5'-UTR and its expression is modulated by G4 resolvase 1”. *Nucleic Acids Res.* 40, 1033-1049 (2012).

338.Vaughn, J. P.; Creacy, S. D.; Routh, E. D.; Joyner-Butt, C.; Jenkins, G. S.; Pauli, S.; Nagamine, Y.; Akman, S. A. “The DEXH protein product of the DHX36 gene is the major

source of tetramolecular quadruplex G4-DNA resolving activity in HeLa cell lysates". *J. Biol. Chem.* 280, 38117-38120 (2005).

339. Harrington, C.; Lan, Y.; Akman, S. A. "The Identification and Characterization of a G4-DNA Resolvase Activity". *J. Biol. Chem.* 272, 24631-24636 (1997).

340. Giri, B.; Smaldino, P. J.; Thys, R. G.; Creacy, S. D.; Routh, E. D.; Hantgan, R. R.; Lattmann, S.; Nagamine, Y.; Akman, S. A.; Vaughn, J. P. "G4 resolvase 1 tightly binds and unwinds unimolecular G4-DNA". *Nucleic Acids Res.* 39, 7161-7178 (2011).

341. Gao, X.; Ma, W.; Nie J.; Zhang, C.; Zhang, J.; Yao, G.; Han, J.; Xu, J.; Hu, B.; Du, Y.; Shi, Q. ; Yang, Z.; Huang, X.; Zhang, Y. "A G-quadruplex DNA structure resolvase, RHAU, is essential for spermatogonia differentiation". *Cell Death Dis.* 6, e1610 (2015).

342. Booy, E. P.; McRae, Ewan K S; McKenna, S. A. "Biochemical characterization of G4 quadruplex telomerase RNA unwinding by the RNA helicase RHAU". *Methods Mol. Biol.* 1259, 125-135 (2015).

343. Sexton, A. N.; Collins, K. "The 5' guanosine tracts of human telomerase RNA are recognized by the G-quadruplex binding domain of the RNA helicase DHX36 and function to increase RNA accumulation". *Mol. Cell. Biol.* 31, 736-743 (2011).

344. Chen, M. C.; Murat, P.; Abecassis, K.; Ferré-D'Amaré, A. R.; Balasubramanian, S. "Insights into the mechanism of a G-quadruplex-unwinding DEAH-box helicase". *Nucleic Acids Res.* 43, 2223-2231 (2015).

345. Meier, M.; Patel, T. R.; Booy, E. P.; Marushchak, O.; Okun, N.; Deo, S.; Howard, R.; McEleney, K.; Harding, S. E.; Stetefeld, J.; McKenna, S. A. "Binding of G-quadruplexes to the N-terminal recognition domain of the RNA helicase associated with AU-rich element (RHAU)". *J. Biol. Chem.* 288, 35014-35027 (2013).

346. Booy, E. P.; Howard, R.; Marushchak, O.; Ariyo, E. O.; Meier, M.; Movakowski, S. K.; Deo, S. R.; Dzananovic, E.; Stetefeld, J.; McKenna, S. A. "The RNA helicase RHAU (DHX36) suppresses expression of the transcription factor PITX1". *Nucleic Acids Res.* 42, 3346-3361 (2014).

347. Gallo, S.; Furler M.; Sigel R. K.O. "*In vitro* Transcription and Purification of RNAs of Different Size". *Chimia* 59, 812-816 (2005).

348. Jenö, P.; Horst, M. "Electroelution of Proteins from Polyacrylamide Gels". In *The Protein Protocols Handbook,* edited by J. M. Walker (Humana Press Inc., New Jersey, USA, 1996).

349. Kao, C.; Rüdisser, S.; Zheng, M. "A simple and efficient method to transcribe RNAs with reduced 3' heterogeneity". *Methods* 23, 201-205 (2001).

350. Hadzic, M. C. A. S.; Kowerko, D.; Börner, R.; König, S. L. B.; Sigel, R. K. O.; Ritter, M. "A Comprehensive Guide for the Analysis of Camera-Based Single-Molecule Fluorescence Experiments". *Submitted.*

351. Mergny, J.-L.; Phan, A.-T.; Lacroix, L. "Following G-quartet formation by UV-spectroscopy". *FEBS Lett.* 435, 74-78 (1998).

352. Mergny, J.-L.; Lacroix, L. "Analysis of Thermal Melting Curves". *Oligonucleotides* 13, 515-537 (2003).

353. Berova, N.; Nakanishi, K.; Woody, R. *Circular dichroism. Principles and applications.* 2nd ed. (Wiley-VCH, New York, 2000).

354. Markley, J. L.; Bax, A.; Arata, Y.; Hilbers, C. W.; Kaptein, R.; Sykes, B. D.; Wright, P. E.; Wüthrich, K. "Recommendations for the Presentation of NMR Structures of Proteins and Nucleic Acids". *J. Mol. Biol.* 280, 933-952 (1998).

355. Cardo, L.; Karunatilaka, K. S.; Rueda, D.; Sigel, R. K. O. "Single molecule FRET characterization of large ribozyme folding". *Methods Mol. Biol.* 848, 227-251 (2012).

356. Mukundan, V. T.; Phan, A. T. “Bulges in G-quadruplexes: broadening the definition of G-quadruplex-forming sequences”. *J. Am. Chem. Soc.* 135, 5017-5028 (2013).

357. Yue, D. J. E.; Lim, K. W.; Phan, A. T. “Formation of (3+1) G-quadruplexes with a long loop by human telomeric DNA spanning five or more repeats”. *J. Am. Chem. Soc.* 133, 11462-11465 (2011).

358. Mergny, J.-L. “Meeting report: Fourth international meeting on G-quadruplex Nucleic Acids (Singapore, July 1-4, 2013)”. *Biochimie* 95, 2320-2325 (2013).

359. Luke, B.; Lingner, J. “TERRA: telomeric repeat-containing RNA”. *EMBO J.* 28**,** 2503-2510 (2009).

360. Morris, M. J.; Wingate, K. L.; Silwal, J.; Leeper, T. C.; Basu, S. “The porphyrin TmPyP4 unfolds the extremely stable G-quadruplex in MT3-MMP mRNA and alleviates its repressive effect to enhance translation in eukaryotic cells”. *Nucleic Acids Res.* 40, 4137-4145 (2012).

361. Shahid, R.; Bugaut, A.; Balasubramanian, S. “The BCL-2 5' untranslated region contains an RNA G-quadruplex-forming motif that modulates protein expression”. *Biochemistry* 49, 8300-8306 (2010).

362. National Center for Biotechnology Information. Genes & Espression. Available at http://www.ncbi.nlm.nih.gov/gene (2015).

363. Page, G.; Lödige, I.; Kögel, D.; Scheidtmann, K. H. “AATF, a novel transcription factor that interacts with Dlk/ZIP kinase and interferes with apoptosis11Accession no. for rat AATF nucleotide sequence at the EMBL GenBank database is RNO238717”. *FEBS Lett.* 462, 187-191 (1999).

364. Neri, A.; Fracchiolla, N. S.; Migliazza, A.; Trecca, D.; Lombardi, L. “The Involvement of the Candidate Proto-Oncogene NFKB2I lyt-10 in Lymphoid Malignancies”. *Leukemia Lymphoma* 23, 43-48 (1996).

365. Abd Karim, N. H.; Mendoza, O.; Shivalingam, A.; Thompson, A. J.; Ghosh, S.; Kuimova, M. K.; Vilar, R. “Salphen metal complexes as tunable G-quadruplex binders and optical probes”. *RSC Adv.* 4, 3355-3363 (2014).

366. Reed, J. E.; Arnal, A. A.; Neidle, S.; Vilar, R. “Stabilization of G-quadruplex DNA and inhibition of telomerase activity by square-planar nickel(II) complexes”. *J. Am. Chem. Soc.* 128, 5992-5993 (2006).

367. Arola-Arnal, A.; Benet-Buchholz, J.; Neidle, S.; Vilar, R. “Effects of metal coordination geometry on stabilization of human telomeric quadruplex DNA by square-planar and square-pyramidal metal complexes”. *Inorg. Chem.* 47, 11910-11919 (2008).

368. Monchaud, D.; Allain, C.; Teulade-Fichou, M.-P. “Development of a fluorescent intercalator displacement assay (G4-FID) for establishing quadruplex-DNA affinity and selectivity of putative ligands”. *Bioorg. Med. Chem. Lett.* 16, 4842-4845 (2006).

369. Mendoza, O.; Abd Karim, N.l H.; Ghosh, S.; Vilar, R. “Metal complexes that target the G-quadruplex structure of the c-Myc oncogene promoter and telomeric DNA”. Poster presented at EUROBIC11, Granada, Spain. Department of Chemistry, Imperial College London, 2012.

370. Pritchard, H. L. PhD thesis. University of Birmingham, 2015.

371. Shamma, M. *The isoquinoline alkaloids. Chemistry and pharmacology* (Academic Press, Inc. London, UK, 1972).

372. Gray, L. T.; Vallur, A. C.; Eddy, J.; Maizels, N. “G quadruplexes are genomewide targets of transcriptional helicases XPB and XPD”. *Nat. Chem. Biol.* 10, 313-318 (2014).

List of figures

List of tables

Appendix

A. Chapter 2

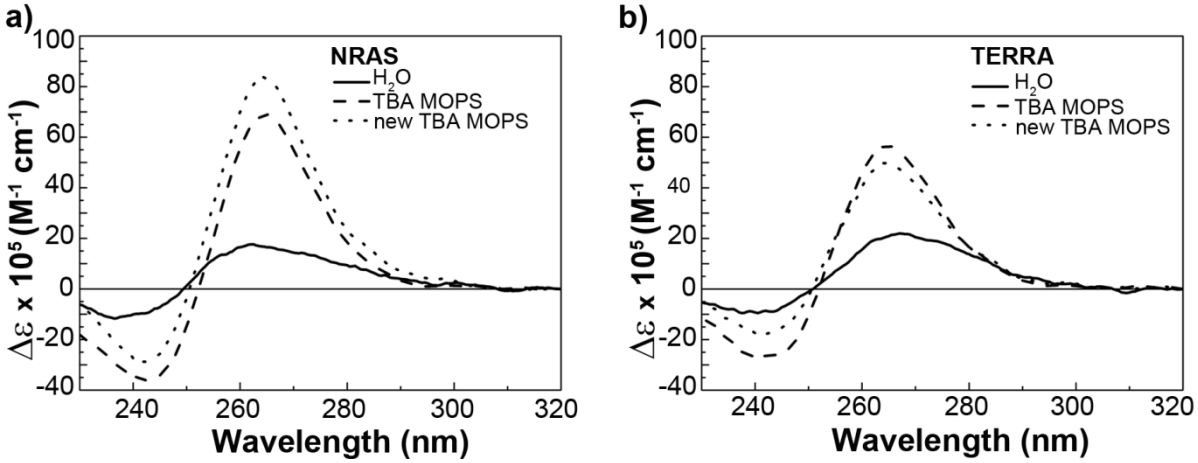


Figure A.1. CD in TBA MOPS. Circular dichroism spectra showing a) NRAS and b) TERRA in 10 mM tetrabutylammonium (TBA) MOPS, pH 7.4.

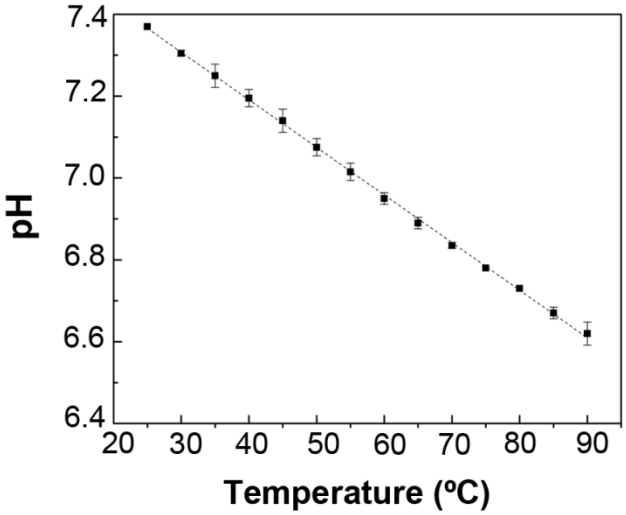


Figure A.2. pH dependence of lithium(I) MOPS buffer. Decrease in pH measured for the 10 mM lithium(I) MOPS buffer, pH 7.4, between room temperature and 90 °C. The errors are calculated as standard deviations between triplicates, and shown as error bars.

Melting curves, measured by CD

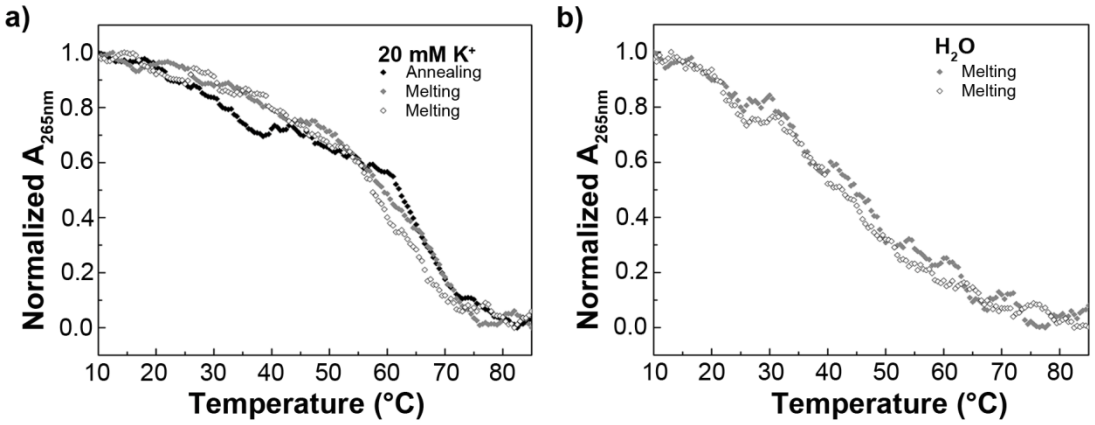


Figure A.3. TERRA CD melting. Melting curves for 4 µM TERRA RNA. a) in 10 mM lithium(I) MOPS, pH 7.4, with 20 mM of K^+; b) in water, without added cations.

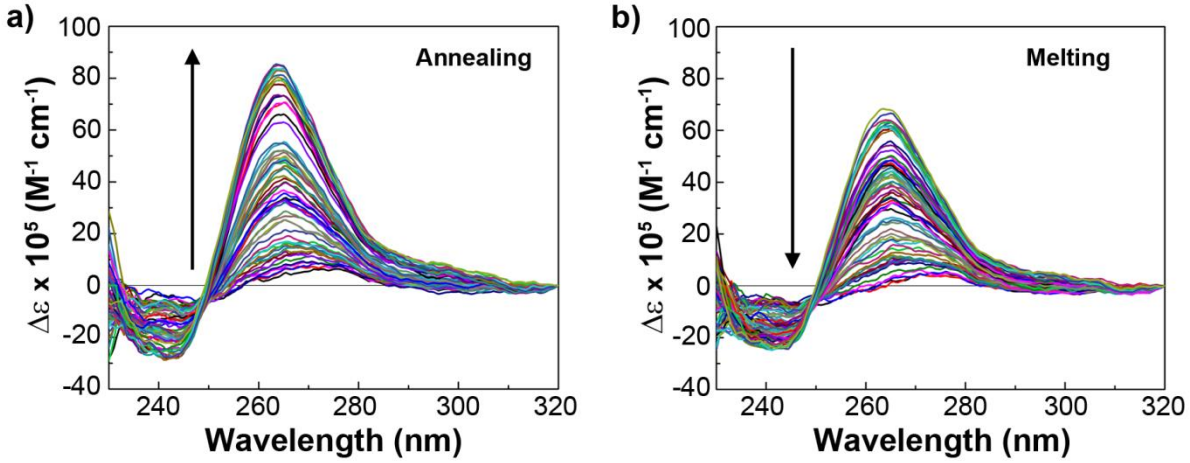


Figure A.4. TERRA CD melting spectra in K^+. CD spectra corresponding to experiments with 4 µM TERRA RNA in 10 mM lithium(I) MOPS, pH 7.4, with 20 mM K^+. a) annealing, 85 to 10 °C and b) melting, 10 to 85 °C.

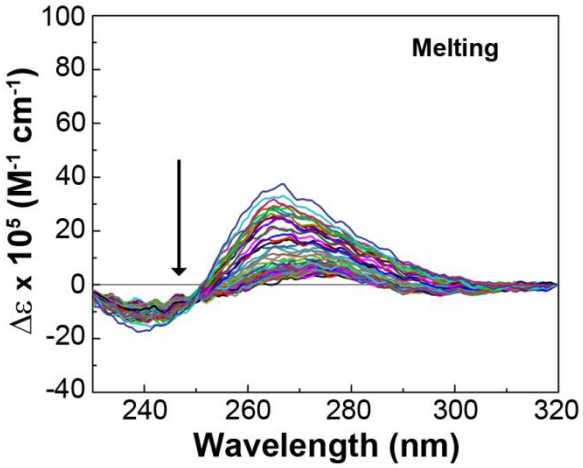


Figure A.5. TERRA CD melting spectra in H_2O. CD spectra corresponding to the melting experiment, from 10 to 85 °C, with 4 µM TERRA RNA in water.

Melting curves, measured by UV

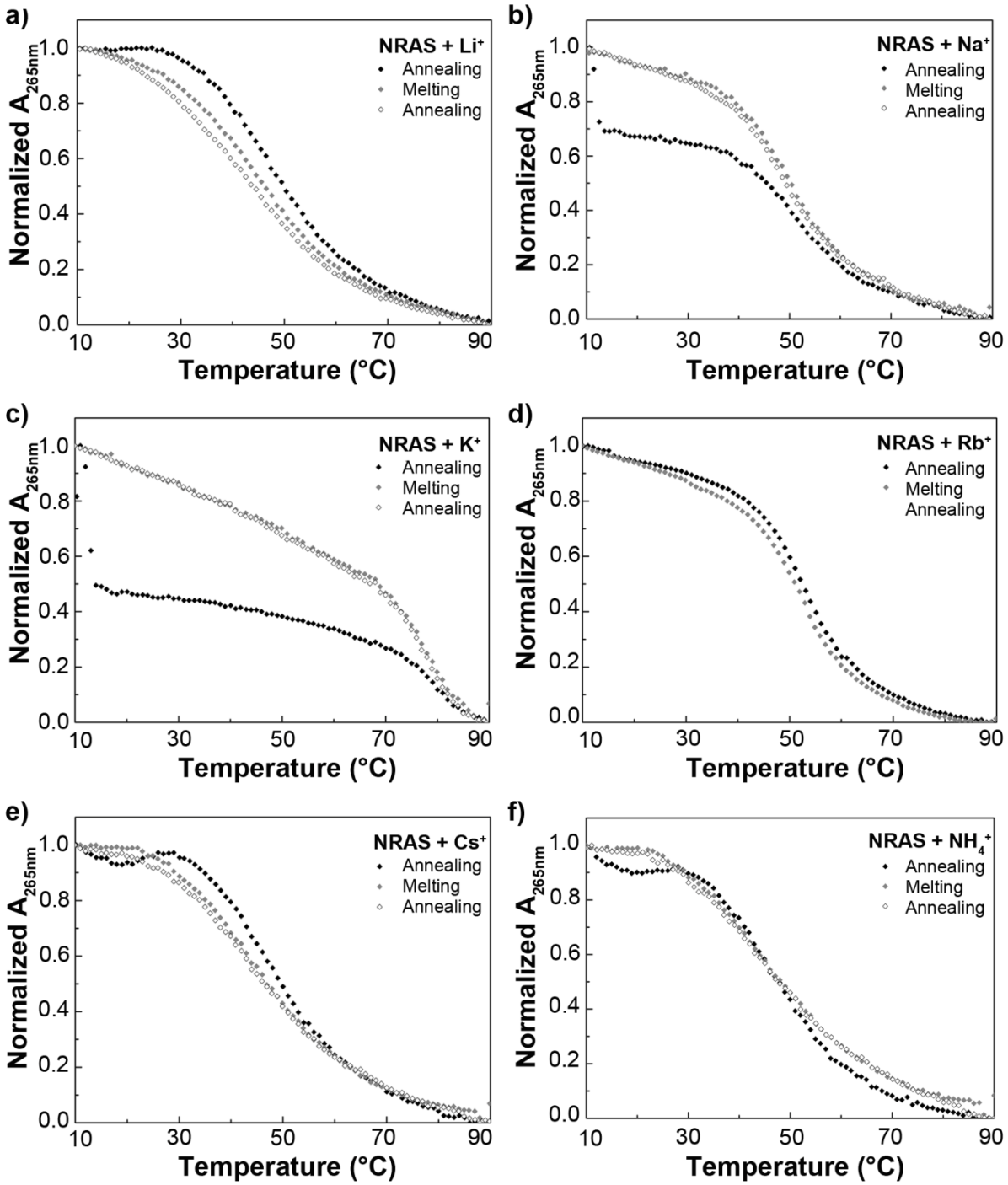


Figure A.6. NRAS with monovalent cations. Melting curves corresponding to 4 µM NRAS RNA in 10 mM lithium(I) MOPS, pH 7.4, with 20 mM of M^+. a) Li^+, b) Na^+, c) K^+, d) Rb^+, e) Cs^+, f) NH_4^+.

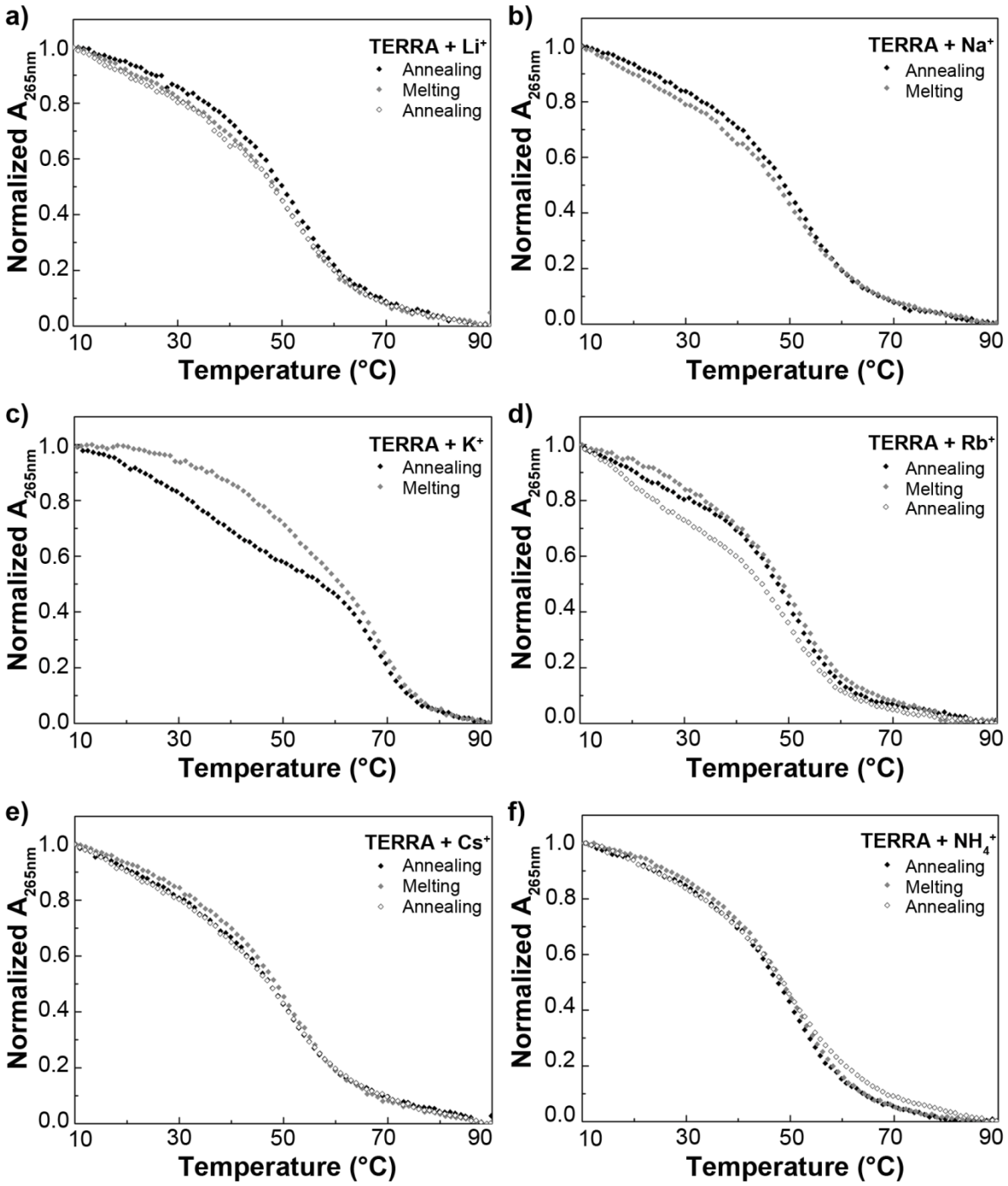


Figure A.7. TERRA with monovalent cations. Melting curves corresponding to 4 µM TERRA RNA in 10 mM lithium(I) MOPS, pH 7.4, with 20 mM of M^+. a) Li^+, b) Na^+, c) K^+, d) Rb^+, e) Cs^+, f) NH_4^+.

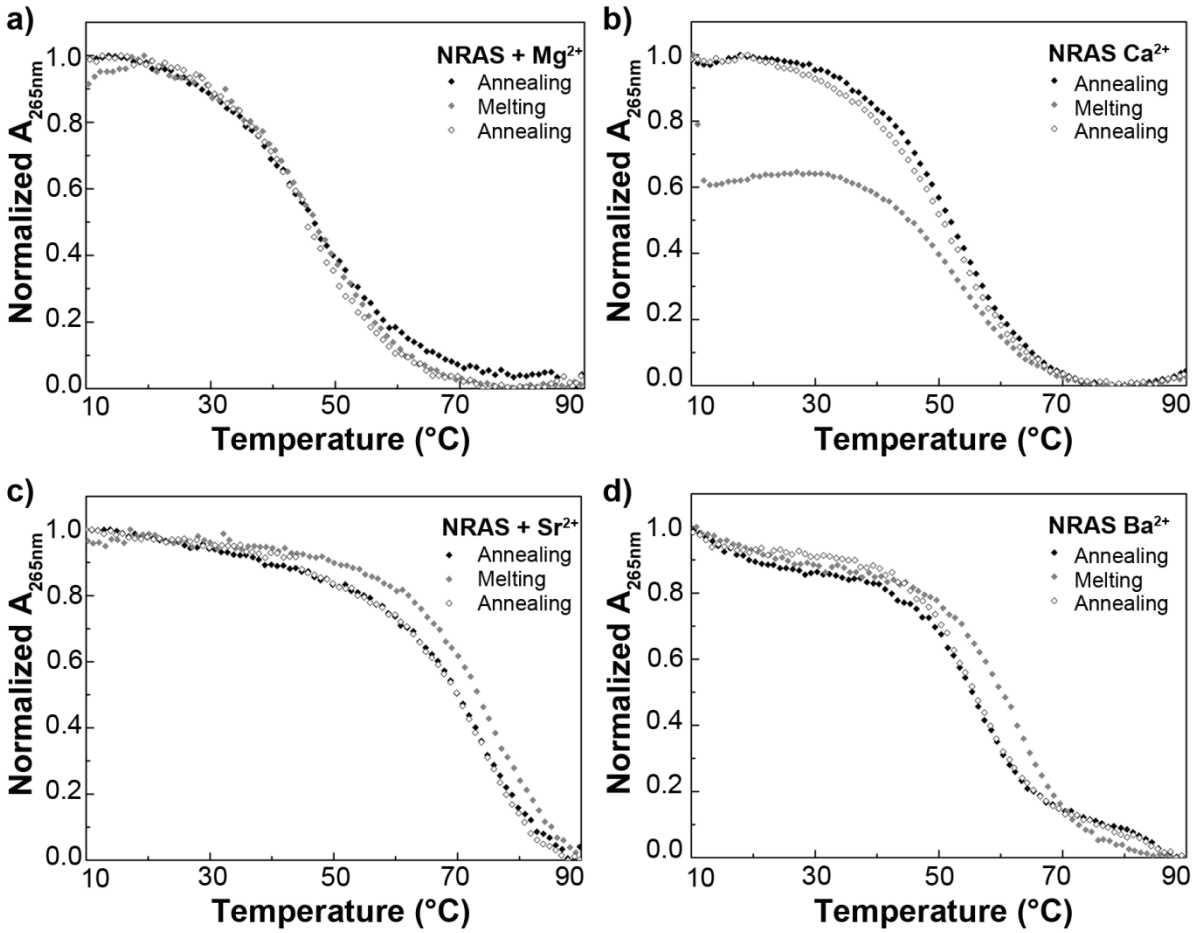


Figure A.8. NRAS with divalent cations. Melting curves corresponding to 4 µM NRAS RNA in 10 mM lithium(I) MOPS, pH 7.4, with 0.1 mM of M^{2+}. a) Mg^{2+}, b) Ca^{2+}, c) Sr^{2+}, d) Ba^{2+}.

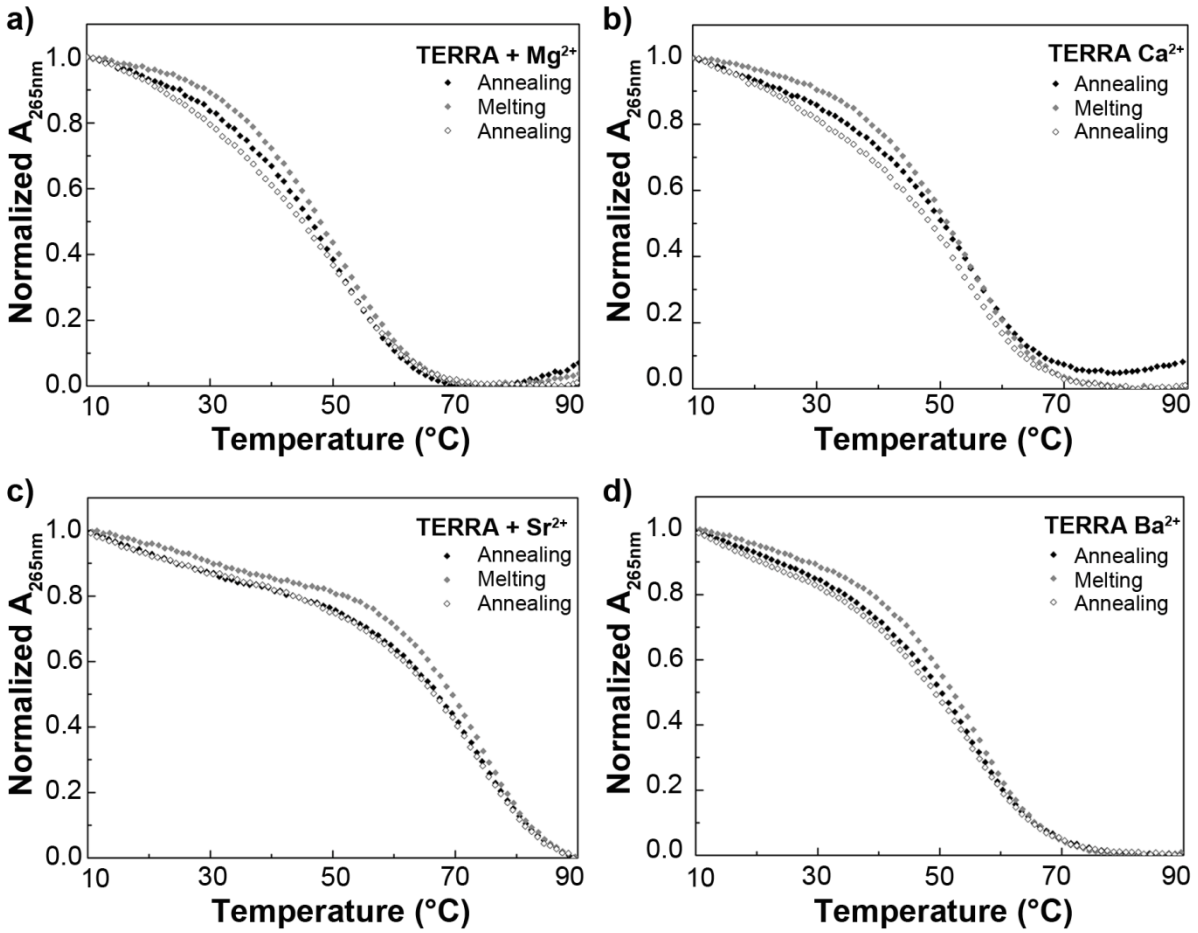


Figure A.9. TERRA with divalent cations. Melting curves corresponding to 4 µM TERRA RNA in 10 mM lithium(I) MOPS, pH 7.4, with 0.1 mM of M^{2+}. a) Mg^{2+}, b) Ca^{2+}, c) Sr^{2+}, d) Ba^{2+}.

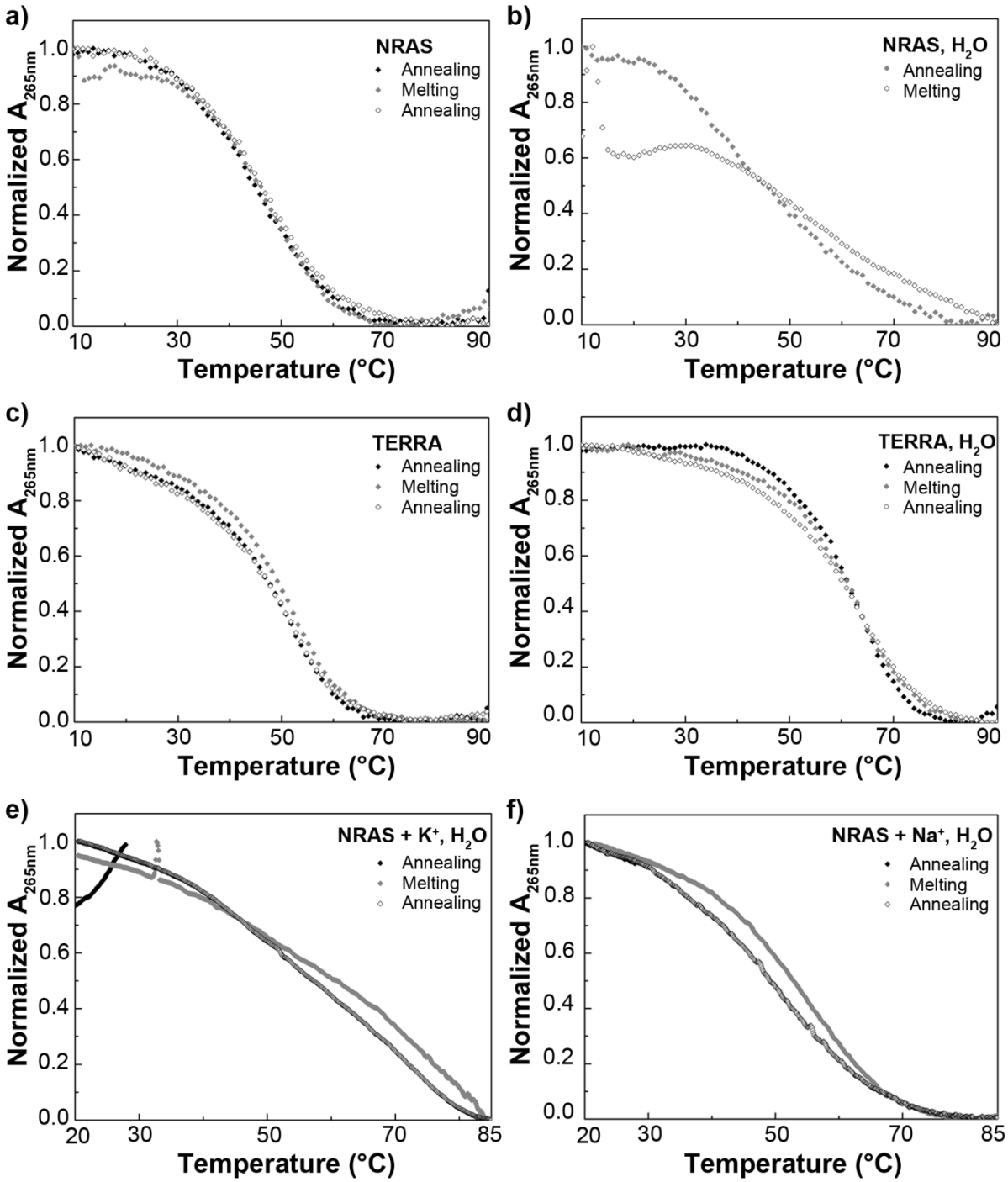


Figure A.10. NRAS and TERRA controls. Melting curves corresponding to 4 µM NRAS and TERRA RNAs. a) and c) in 10 mM lithium(I) MOPS, pH 7.4, without added cations; b) and d) in water, without added cations; e) and f) in water, with 20 mM K^+ or Na^+, respectively.

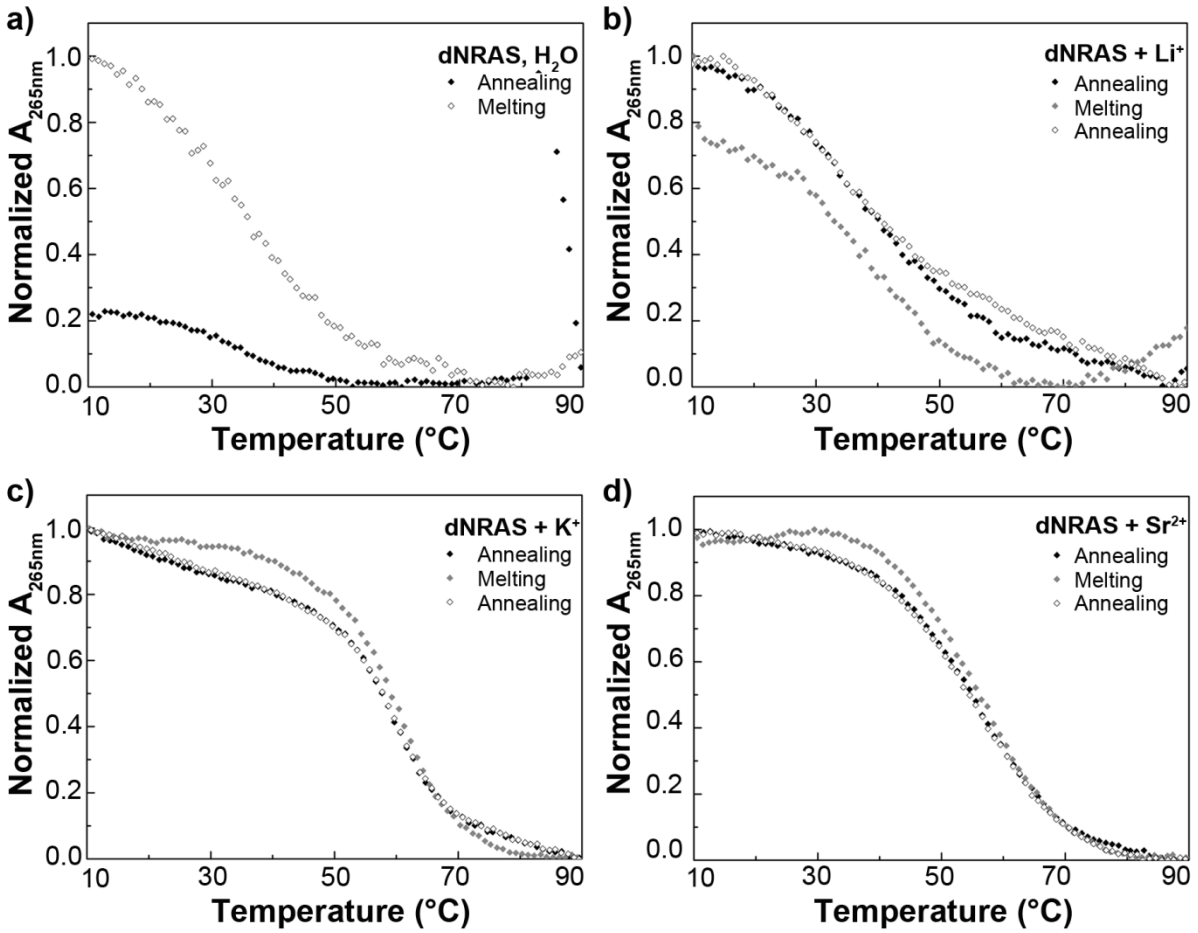


Figure A.11. NRAS DNA. Melting curves corresponding to 4 µM dNRAS DNA in a) water, without added cations; b)-d) 10 mM lithium(I) MOPS, pH 7.4, with 20 mM M^+ or 0.1 mM M^{2+}.

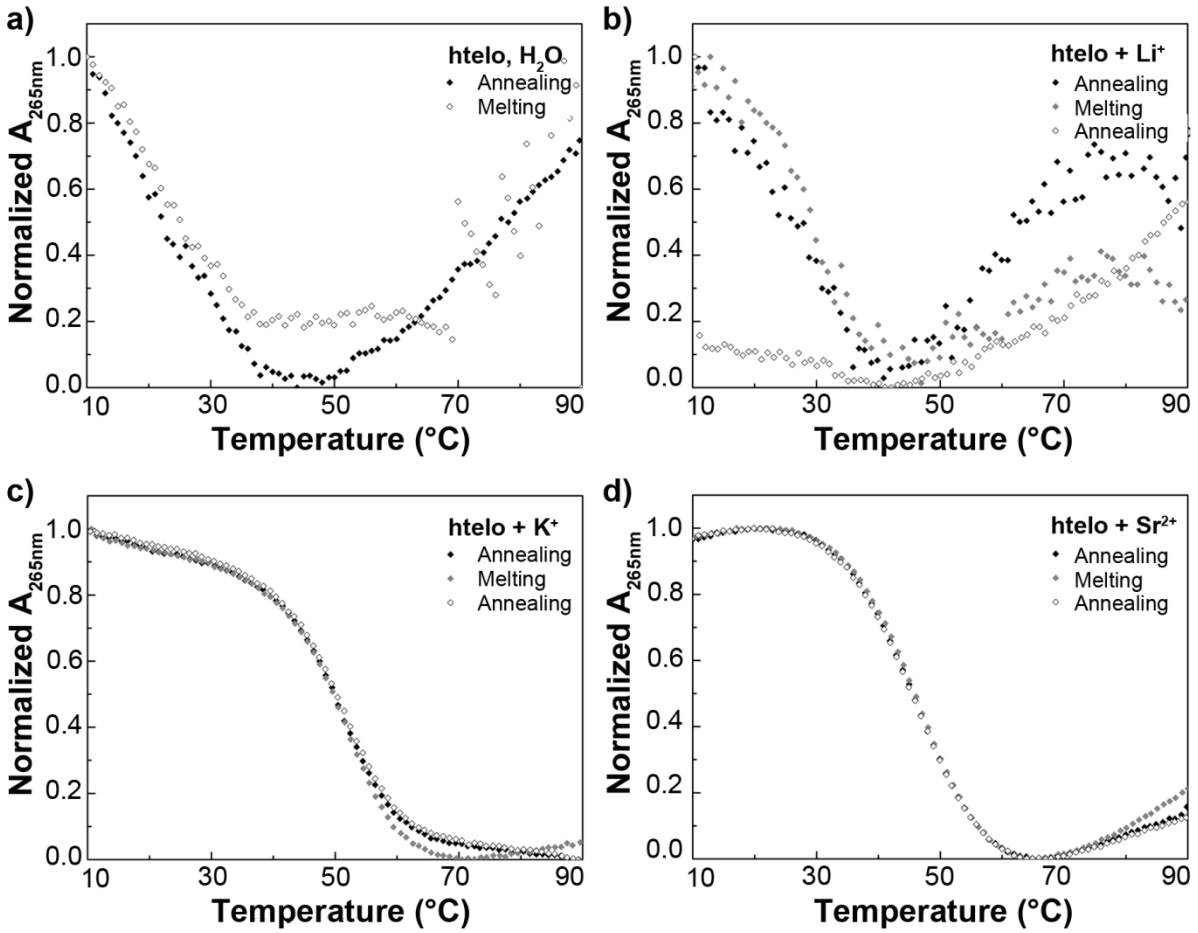


Figure A.12. Telomeric DNA. Melting curves corresponding to 4 µM htelo DNA in a) water, without added cations; b)-d) 10 mM lithium(I) MOPS, pH 7.4, with 20 mM M^+ or 0.1 mM M^{2+}.

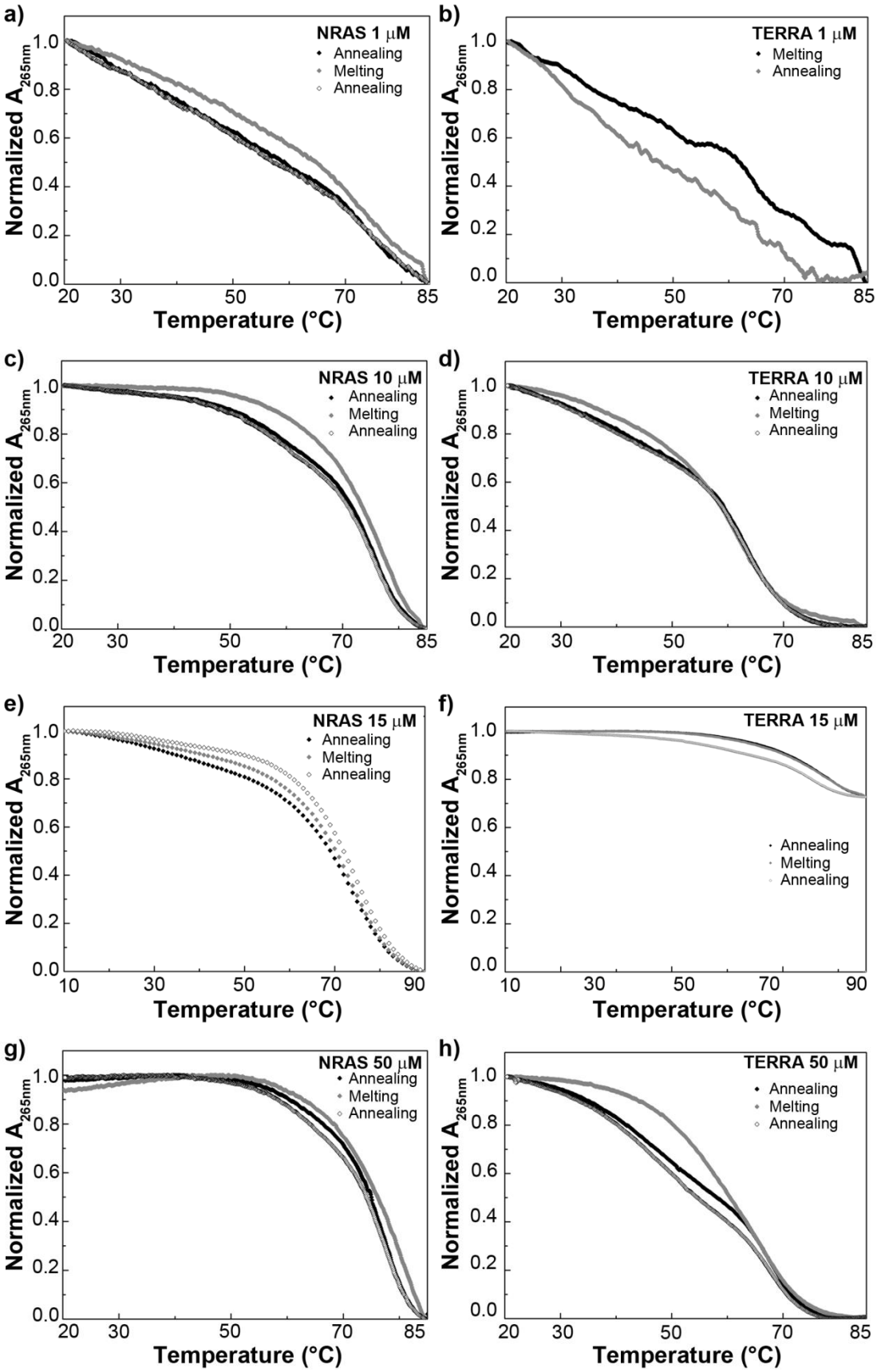


Figure A.13. K^+, RNA concentration. Melting curves corresponding to NRAS or TERRA RNAs 1-50 μM in 10 mM lithium(I) MOPS, pH 7.4, with 20 mM K^+. a) and b) 1 μM, c) and d) 10 μM, e) and f) 15 μM, g) and h) 50 μM.

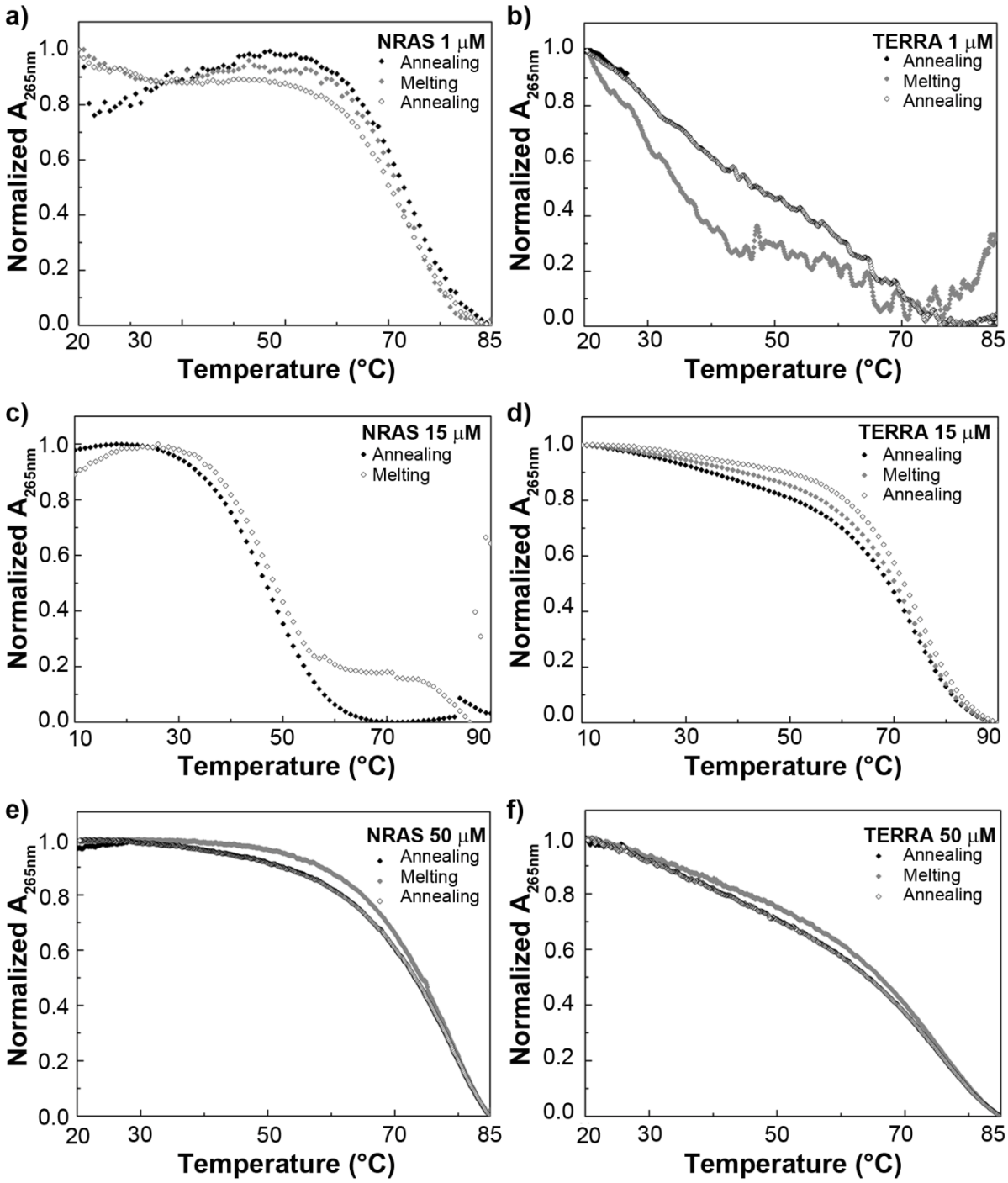
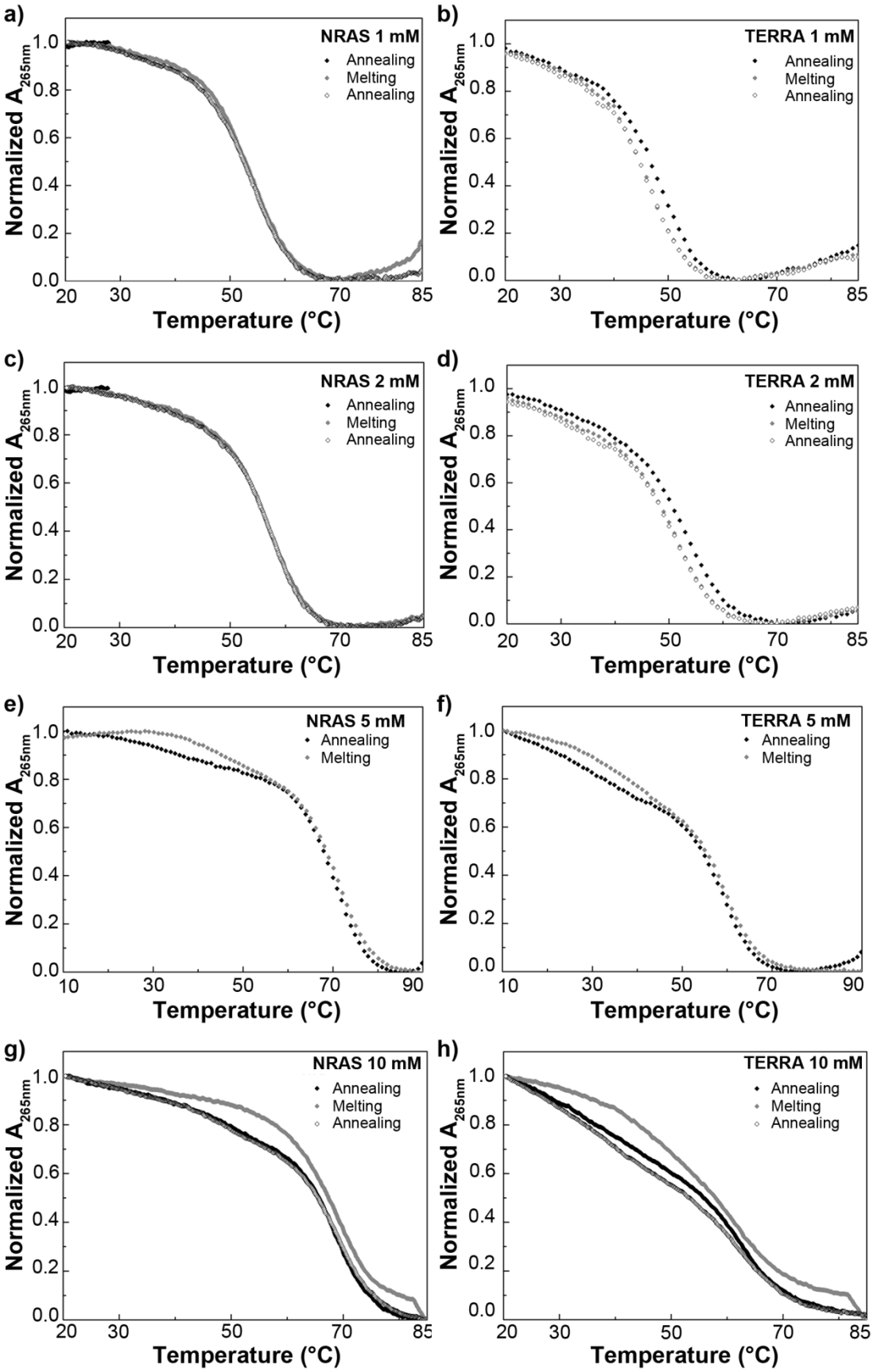
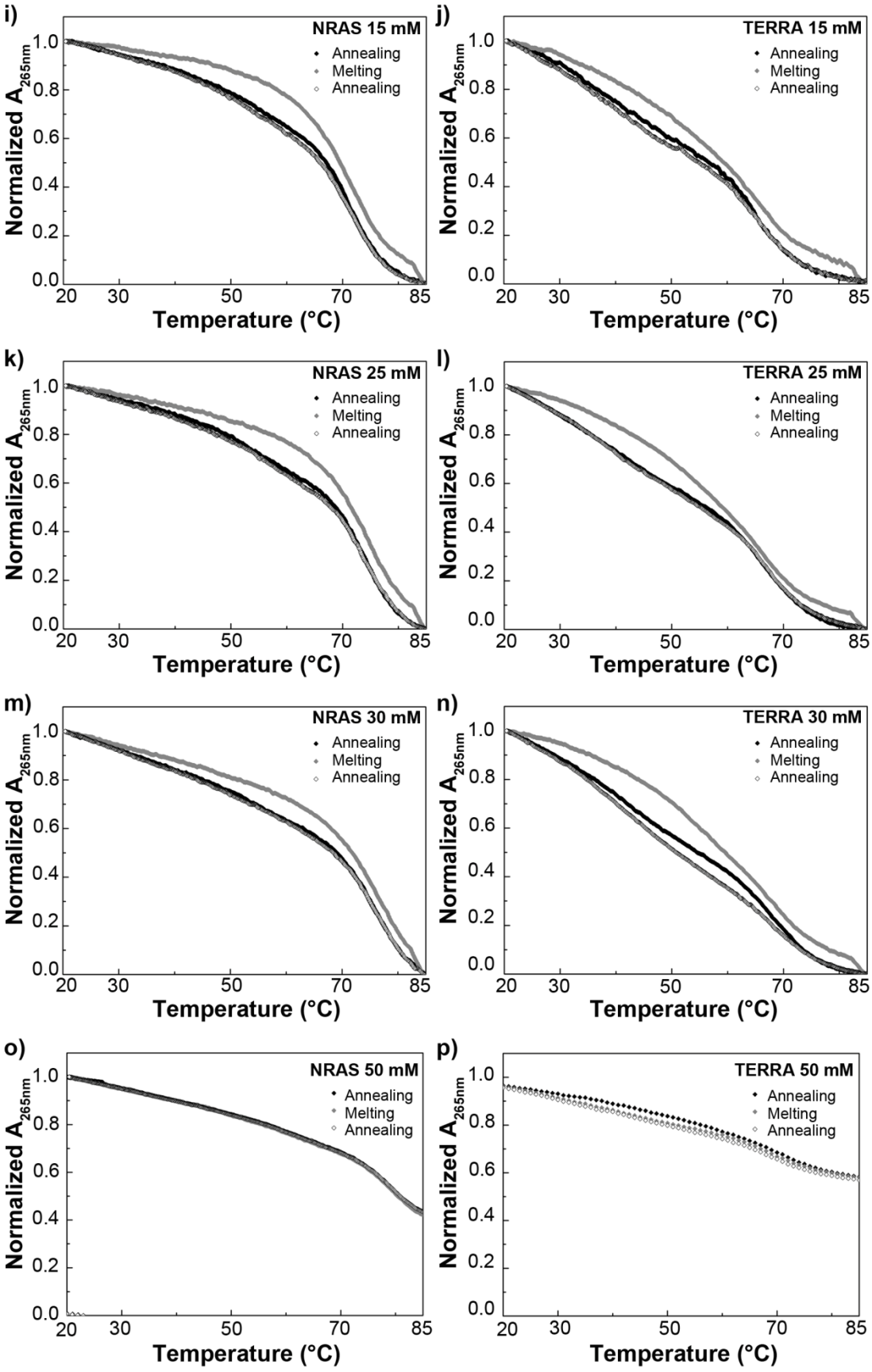


Figure A.14. Sr^{2+}, RNA concentration. Melting curves corresponding to NRAS or TERRA RNAs 1-50 µM in 10 mM lithium(I) MOPS, pH 7.4, with 0.1 mM Sr^{2+}. a) and b) 1 µM, c) and d) 15 µM, e) and f) 50 µM.


a)
NRAS 1 mM
Annealing
Melting
Annealing
Normalized A265nm
Temperature (°C)
b)
TERRA 1 mM
Annealing
Melting
Annealing
Normalized A265nm
Temperature (°C)
c)
NRAS 2 mM
Annealing
Melting
Annealing
Normalized A265nm
Temperature (°C)
d)
TERRA 2 mM
Annealing
Melting
Annealing
Normalized A265nm
Temperature (°C)
e)
NRAS 5 mM
Annealing
Melting
Normalized A265nm
Temperature (°C)
f)
TERRA 5 mM
Annealing
Melting
Normalized A265nm
Temperature (°C)
g)
NRAS 10 mM
Annealing
Melting
Annealing
Normalized A265nm
Temperature (°C)
h)
TERRA 10 mM
Annealing
Melting
Annealing
Normalized A265nm
Temperature (°C)


i)
NRAS 15 mM
j)
TERRA 15 mM
k)
NRAS 25 mM
l)
TERRA 25 mM
m)
NRAS 30 mM
n)
TERRA 30 mM
o)
NRAS 50 mM
p)
TERRA 50 mM
Annealing
Melting
Annealing
Normalized A265nm
Temperature (°C)
0.0
0.2
0.4
0.6
0.8
1.0
20
30
50
70
85

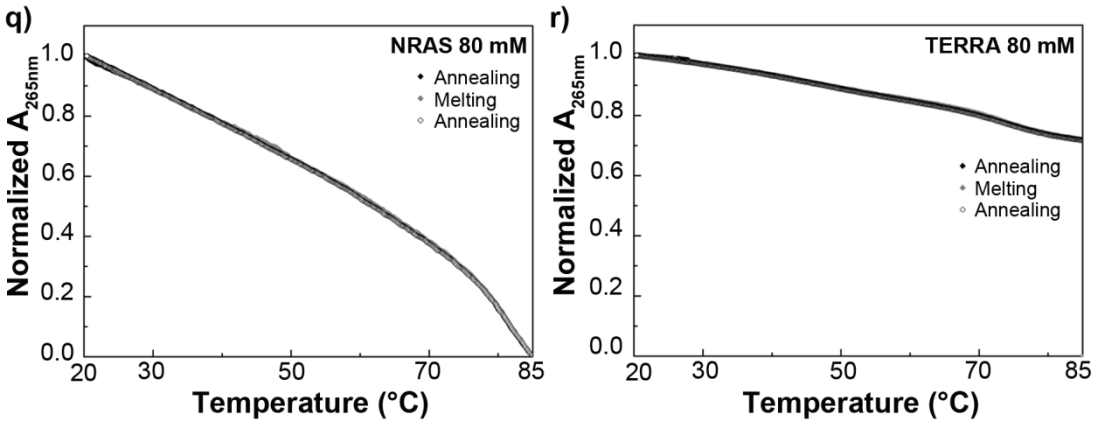
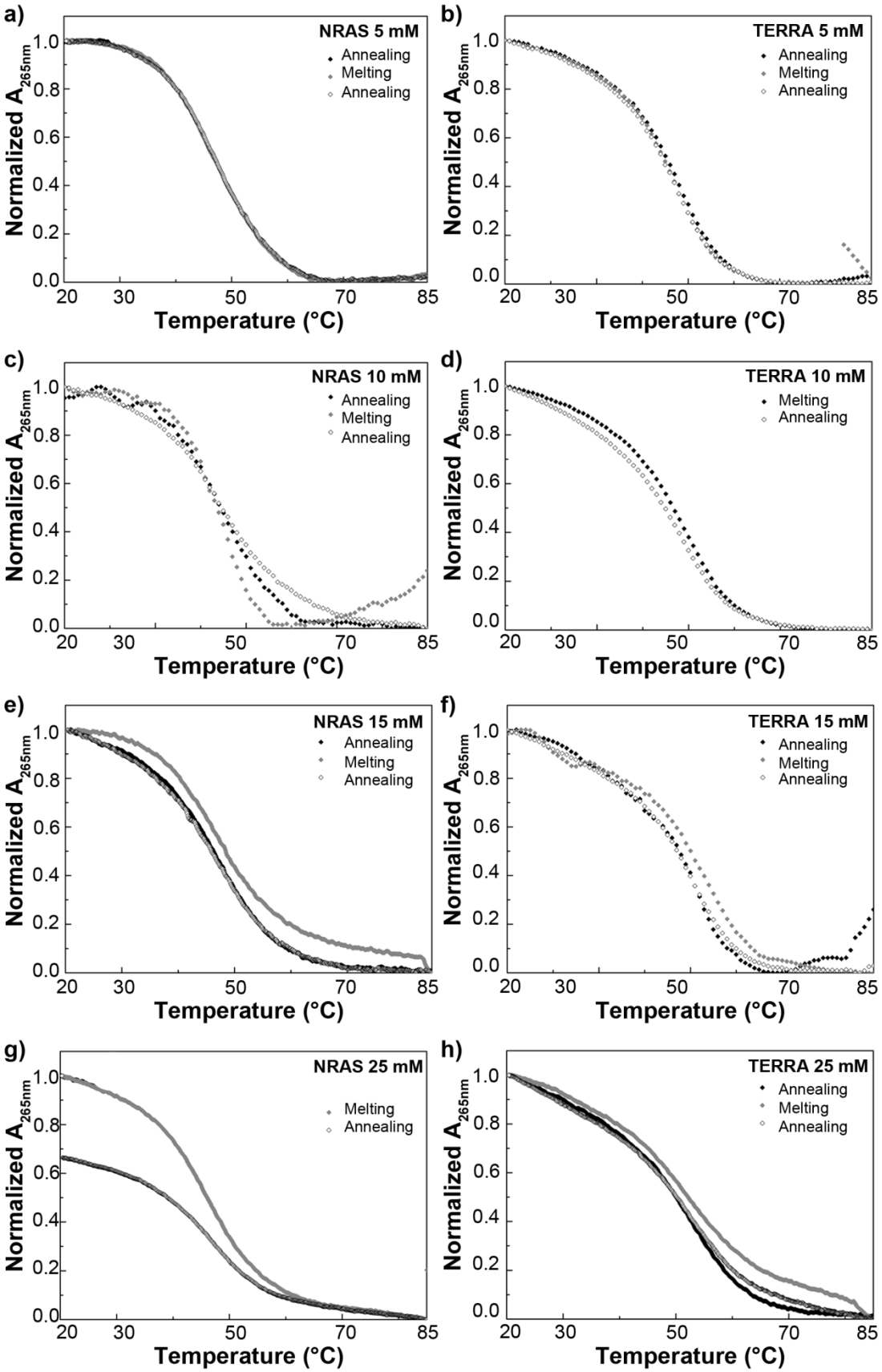


Figure A.15. NRAS and TERRA K^+ titrations followed by UV melting. Melting curves corresponding to 4 µM NRAS or TERRA RNAs in 10 mM lithium(I) MOPS, pH 7.4, upon titration with K^+, 1-80 mM. a) and b) 1 mM, c) and d) 2 mM, e) and f) 5 mM, g) and h) 10 mM, i) and j) 15 mM, k) and l) 25 mM, m) and n) 30 mM, o) and p) 50 mM, q) and r) 80 mM.


a)
NRAS 5 mM
Annealing
Melting
Annealing
b)
TERRA 5 mM
Annealing
Melting
Annealing
c)
NRAS 10 mM
Annealing
Melting
Annealing
d)
TERRA 10 mM
Melting
Annealing
e)
NRAS 15 mM
Annealing
Melting
Annealing
f)
TERRA 15 mM
Annealing
Melting
Annealing
g)
NRAS 25 mM
Melting
Annealing
h)
TERRA 25 mM
Annealing
Melting
Annealing
Normalized A265nm
Temperature (°C)

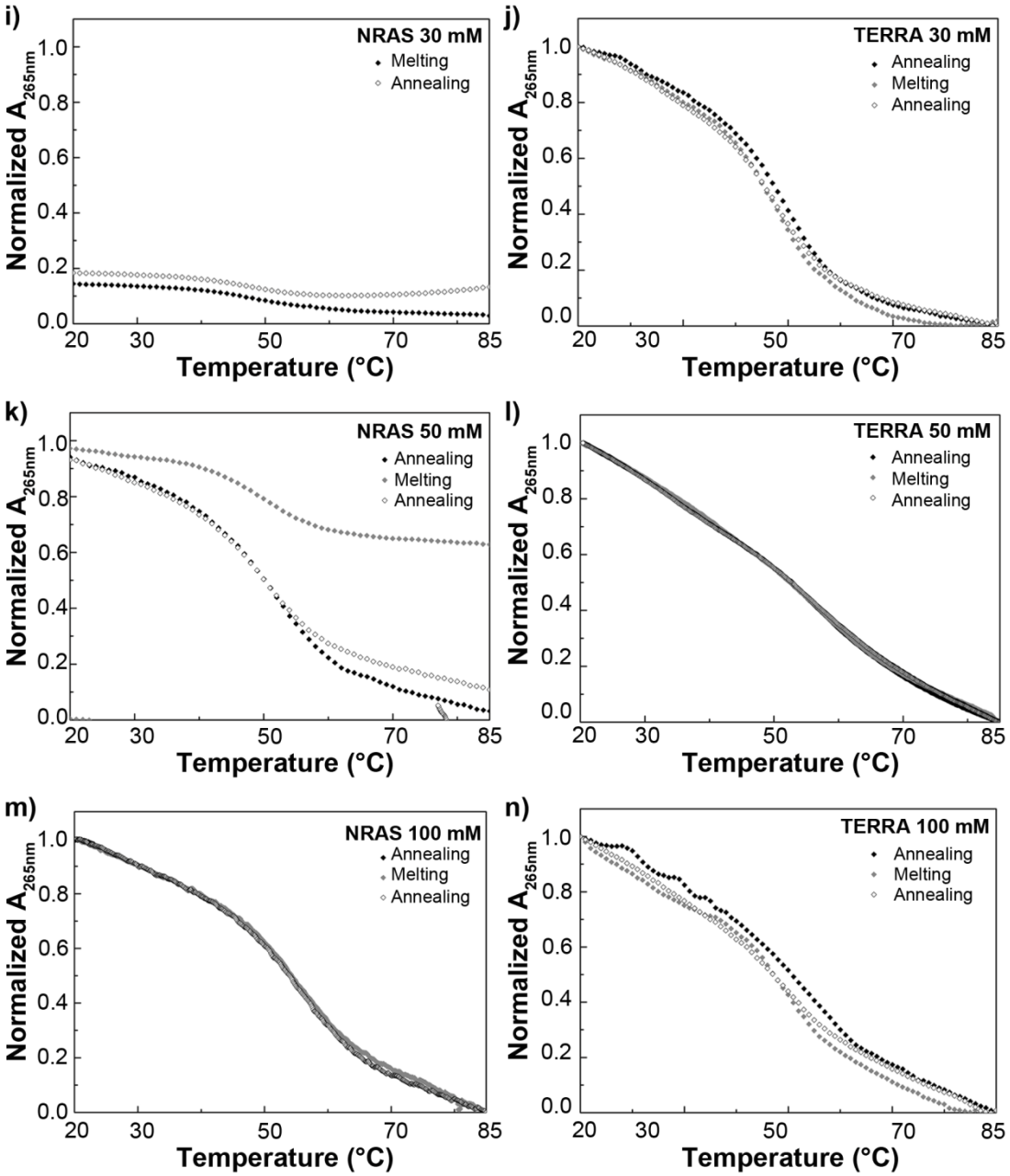


Figure A.16. NRAS and TERRA Na^+ titrations followed by UV melting. Melting curves corresponding to 4 µM NRAS or TERRA RNAs in 10 mM lithium(l) MOPS, pH 7.4, upon titration with Na^+, 5-100 mM. a) and b) 5 mM, c) and d) 10 mM, e) and f) 15 mM, g) and h) 25 mM, i) and j) 30 mM, k) and l) 50 mM, m) and n) 100 mM.

MALDI-MS

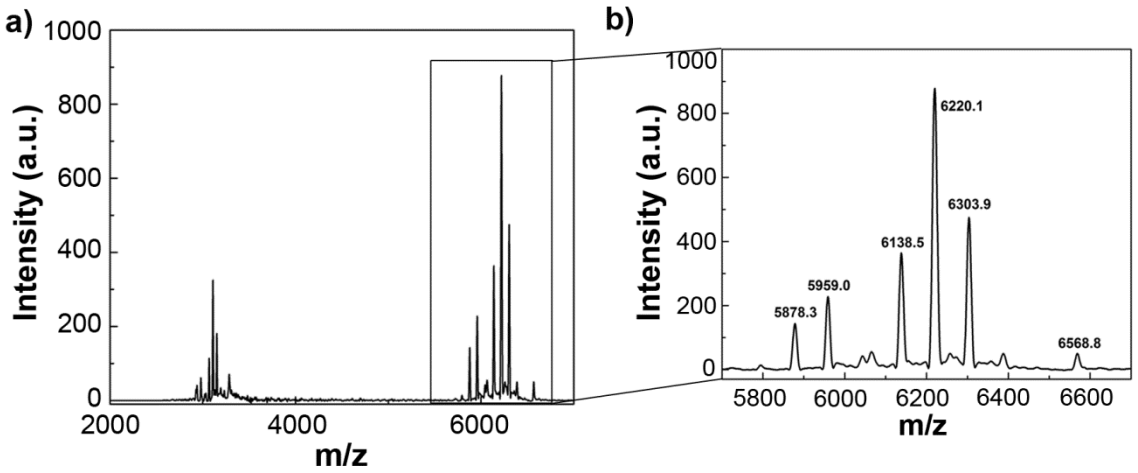


Figure A.17. MALDI-MS spectrum of NRAS with Sr^{2+}. 15 µM NRAS sample with 75 mM Sr^{2+}. a) Molecular peak for the 5'-triphosphate full-length RNA observed at 6220.1 g/mol (calcd. 6217.5 g/mol), b) Zoom-in on the molecular peak.

B. Chapter 3

Native PAGE staining

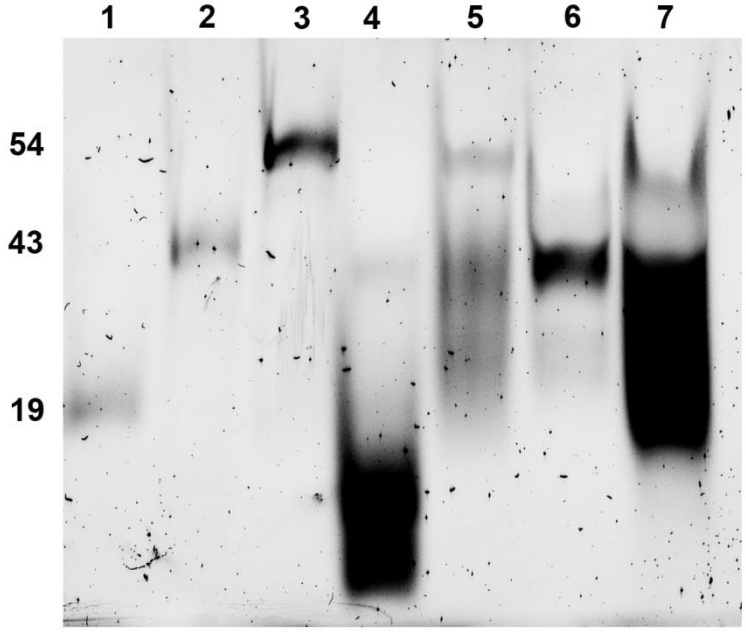


Figure B.1. Staining with GelRed. 15 % native PAGE gel run at 4 °C in 1 × TBE with 20 mM K^+. Short oligonucleotides of known size were used as size markers, (lanes 1-3; 19 nt, 43 nt, and 54 nt). RNA samples: dNRAS (4), NRAS18 (5), tNRAS18 (6), NRAS20 (7).

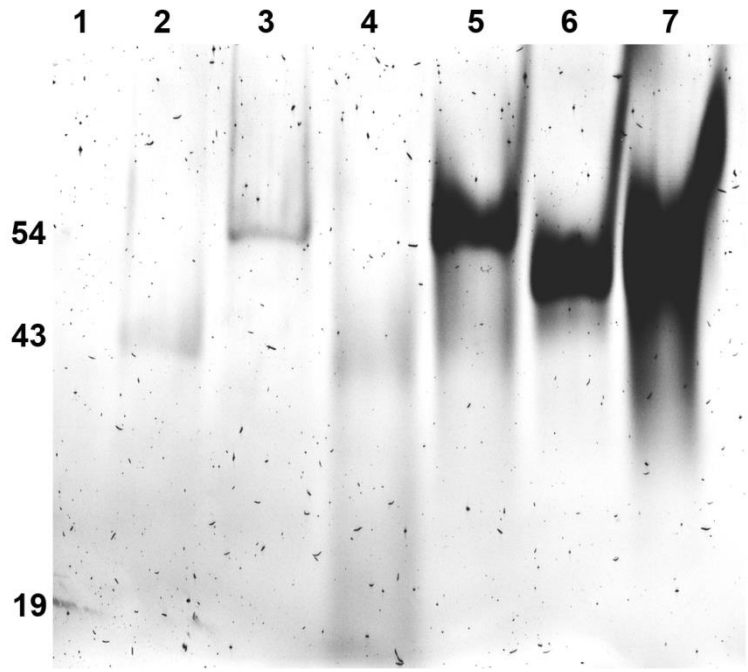


Figure B.2. Staining with EtBr. 15 % native PAGE gel run at 4 °C in 1 × TBE with 20 mM K^+. Short oligonucleotides of known size were used as size markers (lanes 1-3; 19 nt, 43 nt, and 54 nt). RNA samples: dNRAS (4), NRAS18 (5), tNRAS18 (6), NRAS20 (7).

Modified sequences form G4s

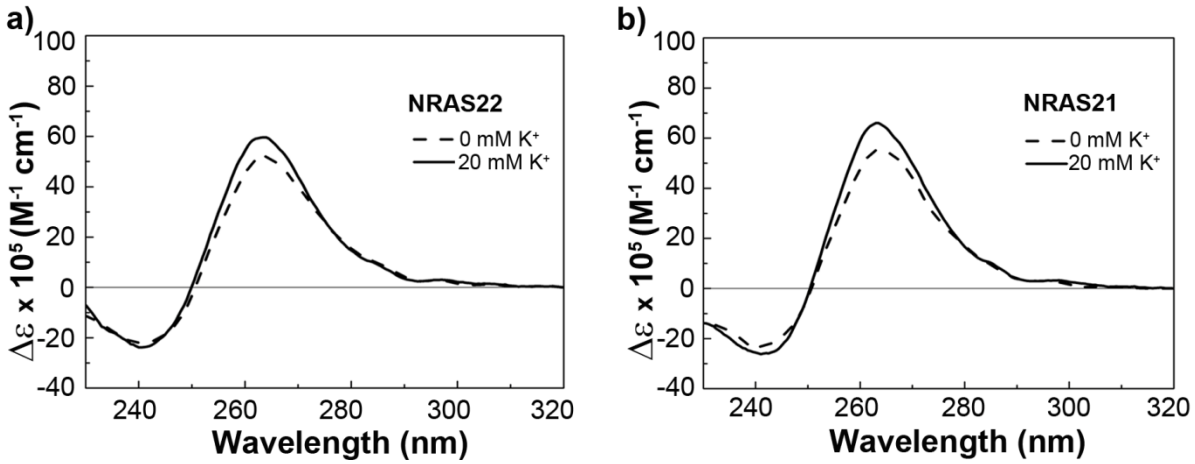


Figure B.3. Modified NRAS folds as a G4. Circular dichroism showing parallel G4 formation in the modified NRAS RNA sequences, a) NRAS22 and b) NRAS21. Measurement conditions: 10 µM RNA in 10 mM lithium(I) MOPS pH 7.4, with (full line) or without (dotted line) 20 mM K^+.

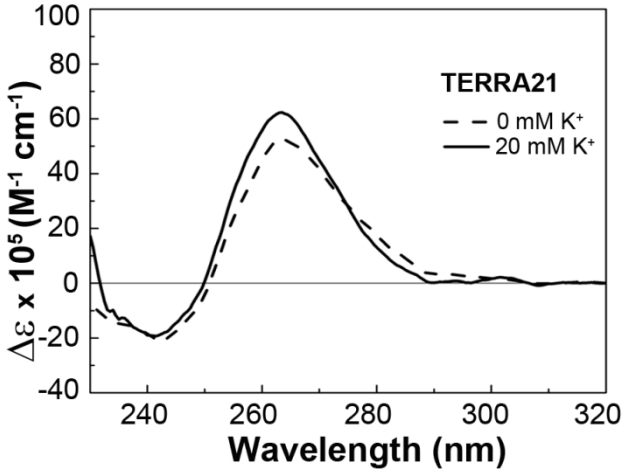


Figure B.4. Modified TERRA folds as a G4. Circular dichroism spectrum showing parallel G4 formation in the modified TERRA RNA sequence, TERRA21. Measurement conditions: 10 µM RNA in 10 mM lithium(I) MOPS pH 7.4, with (full line) or without (dotted line) 20 mM K^+.

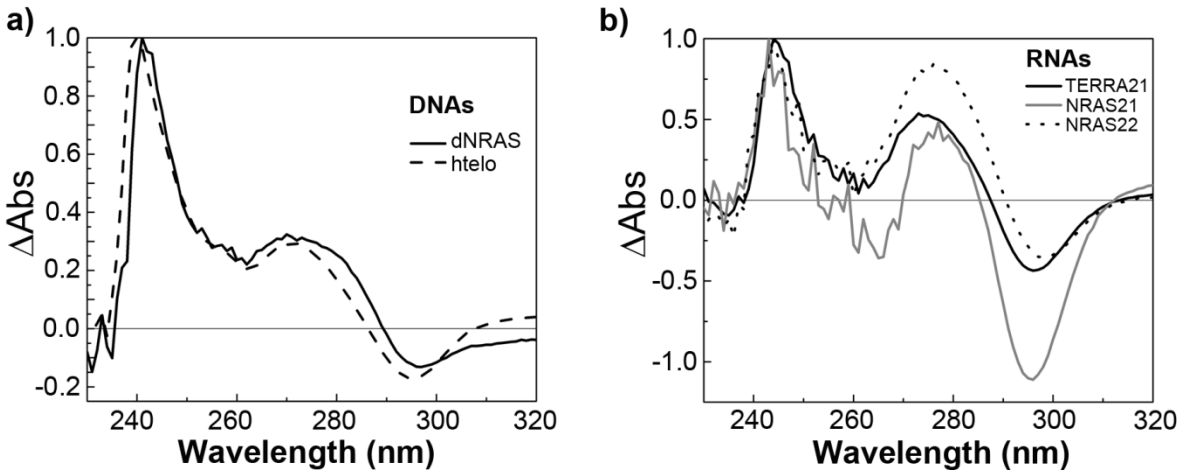


Figure B.5. The corresponding DNAs and the modified RNAs fold into G4s. Thermal difference spectra providing additional proof that a) the DNA sequences, dNRAS (full line) and htelo (dotted line) and b) the modified RNA sequences, TERRA21 (full line), NRAS21 (grey line) and NRAS22 (dotted line), are all folding into G4 structures. Measurement conditions: 10 µM RNA in 10 mM lithium(I) MOPS pH 7.4, with 20 mM K^+.

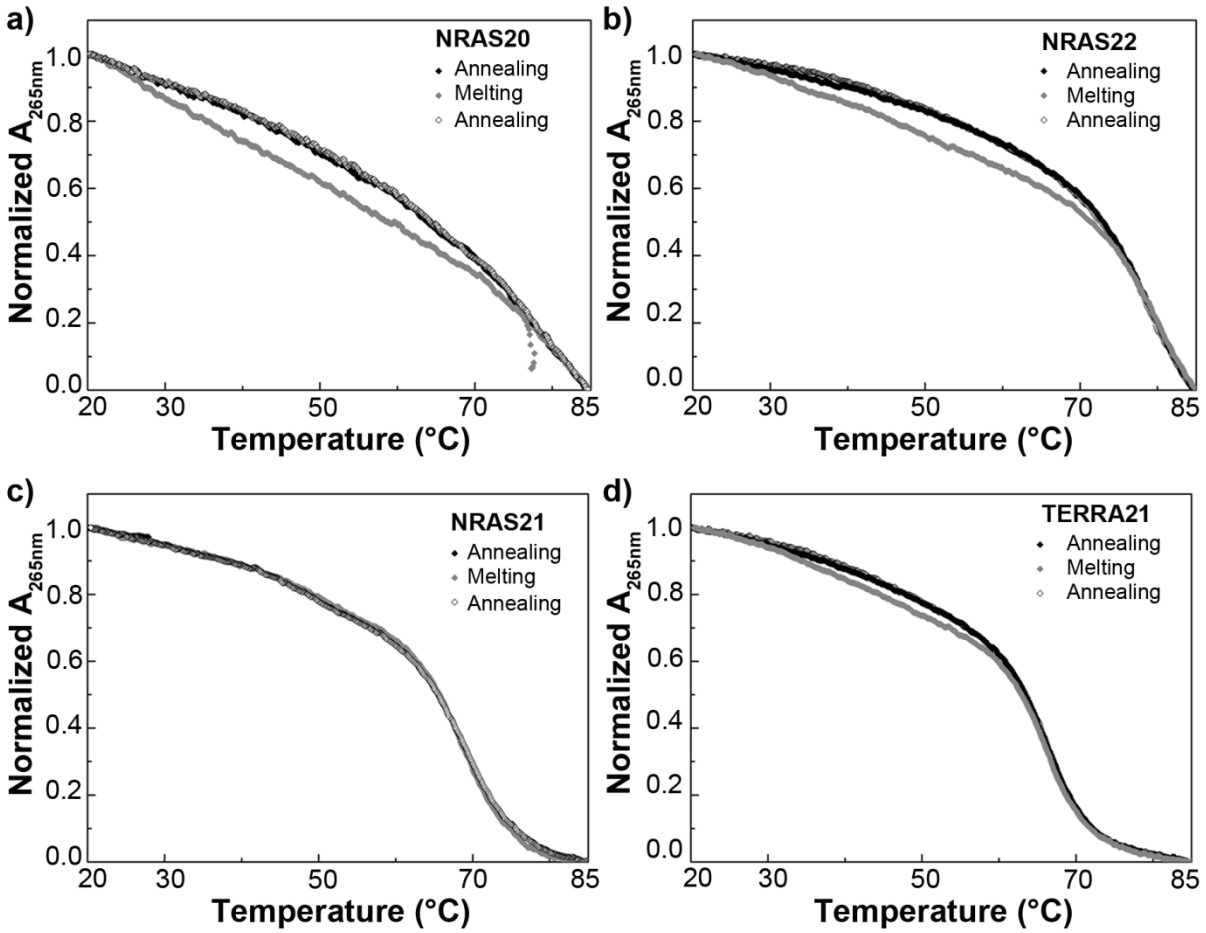


Figure B.6. NRAS and TERRA modified sequences. Melting curves of the modified RNA sequences, measured in 10 mM lithium(I) MOPS, pH 7.4, with 20 mM K^+. a) NRAS20, b) NRAS22, c) NRAS21, and d) TERRA21.

C. Chapter 4

Analytical HPLC

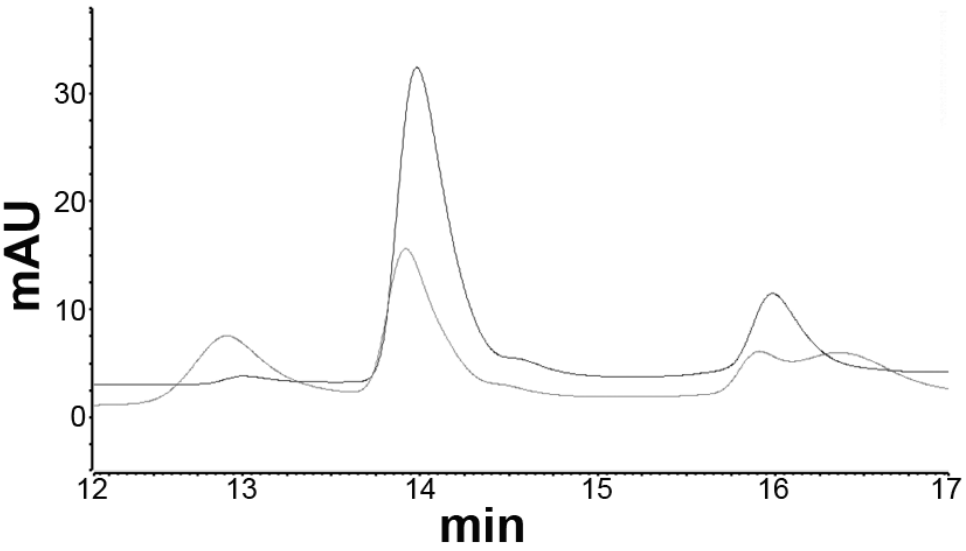


Figure C.1. SCy5-NRAS before HPCL purification. HPLC chromatogram of the reaction mixture (in gray) after SCy5 relabeling. Detection at 260 nm, using a SCy5 solution (in black) as control.

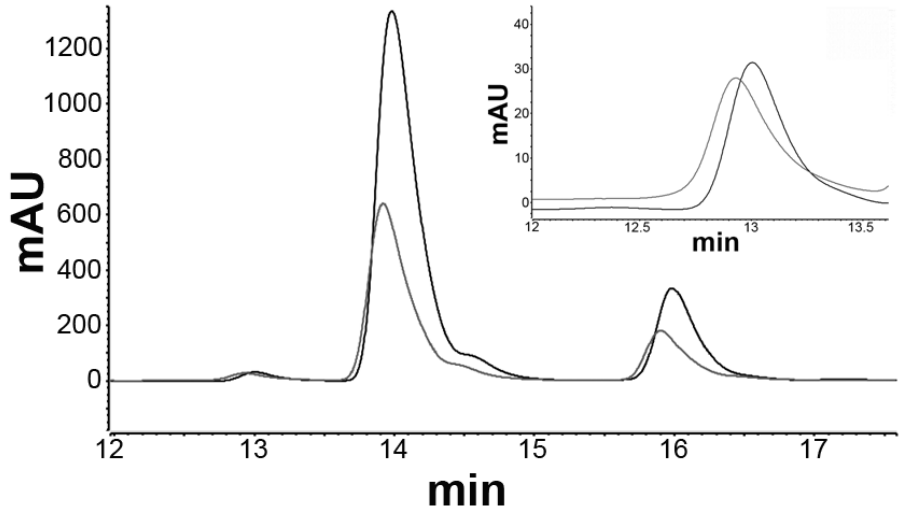


Figure C.2. SCy5-NRAS before HPLC purification. HPLC chromatogram of the reaction mixture (in gray) after SCy5 relabeling. Detection at 646 nm, using a SCy5 solution (in black) as control. The insert shows the enlarged peak at ~ 13 min, corresponding to the labeled NRAS RNA.

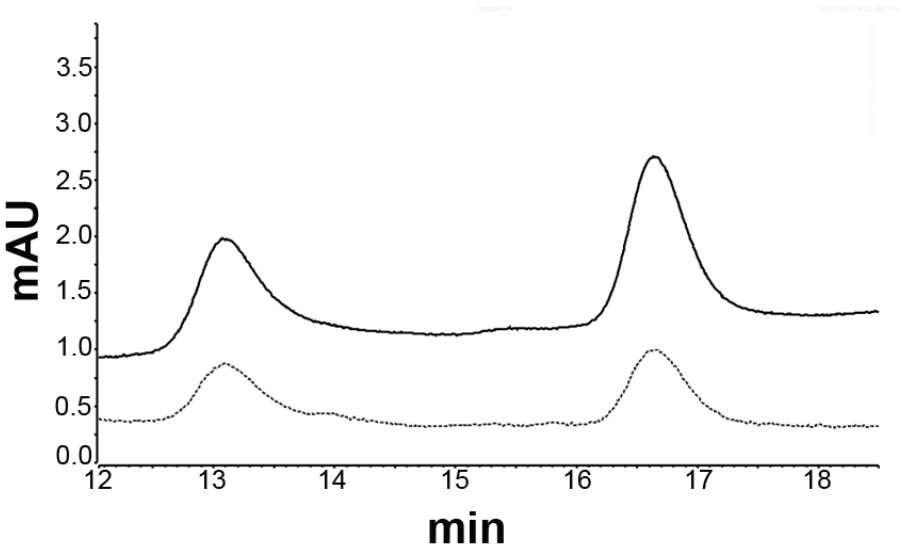


Figure C.3. HPLC-purified SCy5-NRAS. HPLC chromatogram of the 5'-labeled, HPLC-purified SCy5-NRAS RNA, detected at 260 nm (full line) or at 646 nm (dotted line).

Denaturing, analytical, fluorescent PAGE

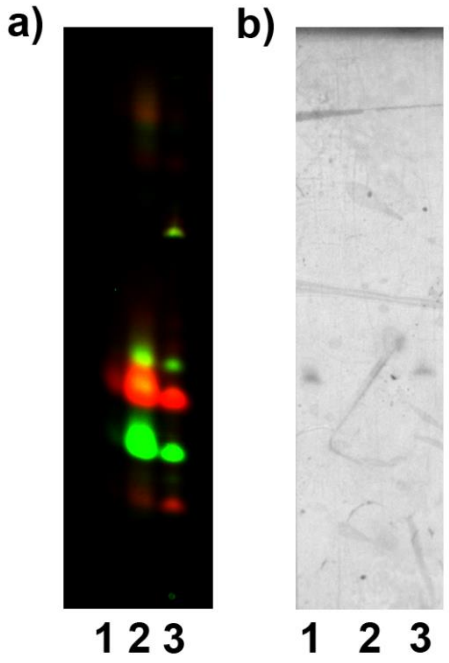


Figure C.4. Labeled SCy5-NRAS analyzed by PAGE. 10 % denaturing PAGE showing the presence of SCy5-labeled NRAS RNA. Lanes: 1) NRAS RNA, 2) SCy5, 3) reaction mixture (labeled NRAS). a) Detection by fluorescent scanning at 532 and 646 nm. Both fluorescent signals (green and red, respectively) are superimposed. b) Detection by UV shadowing. Lane 3, loaded with the labeling reaction mixture, is seen by both the fluorescent and UV detections, and contains both the SCy5 and the RNA.

Labeling efficiency measured by UV-vis.

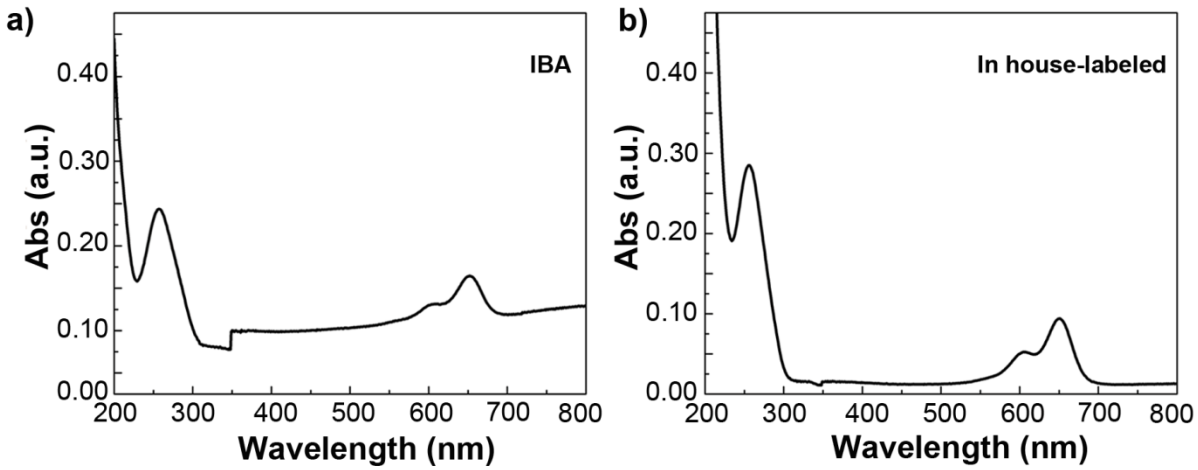


Figure C.5. Absorbance spectra of SCy5-NRAS from IBA and SCy5-NRAS labeled in-house. UV-vis spectra between 200 and 800 nm used to determine the labeling efficiency in a) the 5'-labeled SCy5-NRAS RNA commercially obtained from IBA and b) the 5'-labeled SCy5-NRAS RNA prepared in-house.

Circular dichroism of the smFRET construct

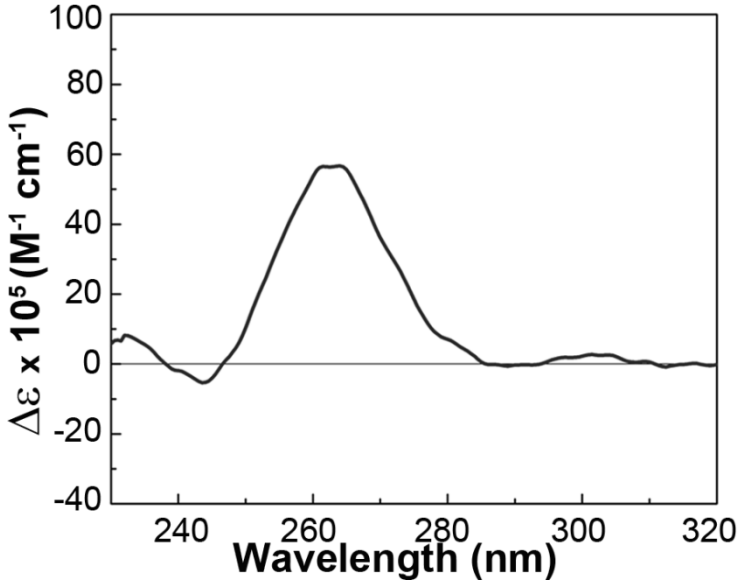


Figure C.6. G4 formation in the NRAS smFRET construct. Circular dichroism spectrum confirming the folding into a G4 for the NRAS construct used in smFRET. Measurement conditions: 10 mM lithium(I) MOPS, pH 7.4, with 20 mM $KClO_4$. A blank measured with the unlabeled 21-nt RNA:DNA duplex corresponding to the smFRET construct was subtracted from the SCy5-NRAS:DNA.

Melting curves

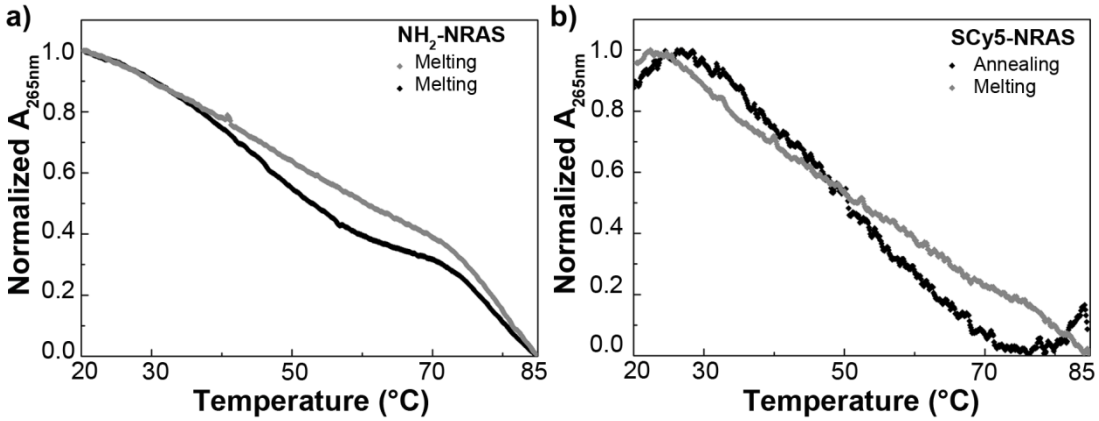


Figure C.7. NRAS for smFRET. Melting curves corresponding to a) the amino-modified 43-nt NRAS sequence and b) the SCy5-labeled 43-nt NRAS used in smFRET experiments. Both RNAs were measured at 4 µM in 10 mM lithium(I) MOPS, pH 7.4, with 20 mM K^+.

smFRET

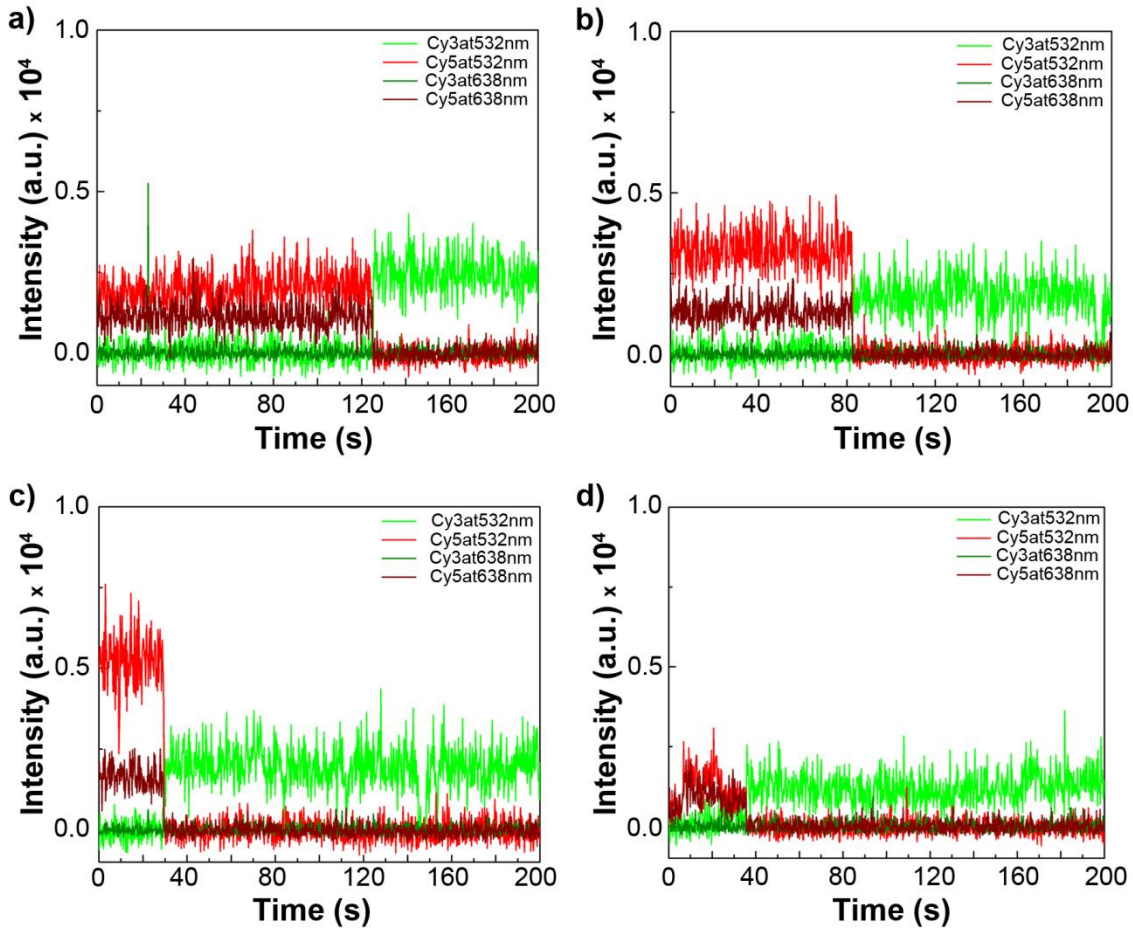
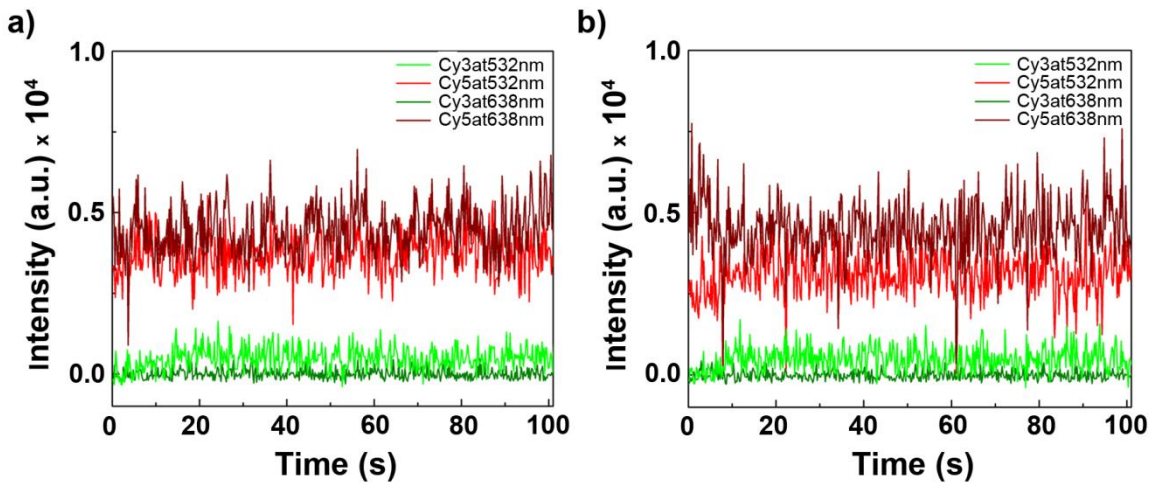


Figure C.8. Further examples of anticorrelation. Traces of fluorescence intensity vs. time obtained in 50 mM lithium(I) MOPS, pH 7.4, with ALEX, at 50 pM RNA and 5 pM DNA (1:0.1 ratio) without added K^+. All traces show anticorrelation upon SCy5 bleaching.



Traces of bi-labeled molecules without anticorrelation. Traces of fluorescence intensity vs. time obtained in 50 mM lithium(I) MOPS, pH 7.4, with ALEX, at 50 pM RNA and 100 pM DNA (1:2 ratio) with 100 mM K^+.

Curriculum Vitae

Personal data

Name	Helena Guiset Miserachs	Date of Birth	29/10/1986
Place of Birth	Barcelona, Spain	Nationality	Spanish

Education

2004-2009 B. Sc. studies in Chemistry at the University of Barcelona, Spain

2009-2011 M. Sc. Studies in Advanced Chemistry, Specialization in Inorganic Chemistry, at the University of Barcelona, Spain

Master project: "Novel ruthenium(II)-arene compounds with intercalating ligands, functionalized adenines and ligands inhibiting glutathione-*S*-transferase with potential antitumor activity" in the group of Prof. Virtudes Moreno

2012-2015 Ph.D. studies in the group of Prof. Roland Sigel, University of Zurich, Switzerland; part of the Chemical and Molecular Sciences Zurich (CMSZH) Graduate School

Practical Experience and Teaching

2007 one-month internship in the Toxicology Department, General Lab Barcelona, Spain – clinical analysis of drugs in biological samples

2008-2009 undergraduate research project, Department of Chemistry, Imperial College London, Dr. Robert P. Davies laboratory – "Novel sulfur based ligands for biomimetic chemistry"

2009 three-month scholarship internship in the Bioinorganic Chemistry group, Prof. Virtudes Moreno, University of Barcelona, Spain – synthesis and biological studies of novel platinum antitumor complexes

2010 teaching assistant in the exercises of the basic chemistry course "Chemical equilibria and reactivity" at the EPF Lausanne, Switzerland

2010 research stay, Prof. Paul J. Dyson, EPF Lausanne, Switzerland – "Organometallic ruthenium compounds inhibiting glutathione-*S*-transferase as potential anticancer drugs"

2011 research apprentice, Dr. John J. Rossi's laboratory, Department of Molecular and Cellular Biology, City of Hope Comprehensive Cancer Center, Duarte, CA, USA – RNA interference for the treatment of mantle cell lymphoma

2012-2013 teaching assistant at the 3-week laboratory course for synthetic inorganic chemistry for third-year Chemistry students, University of Zurich

2013 5-week research stay at Imperial College London, Department of Chemistry, Prof. R. Vilar's laboratory

2013-2014 teaching assistant for first and second-semester basic chemistry exercise session, University of Zurich

Awards and Scholarships

2008-2009 Erasmus Scholarship, 6 months, European Union

2009 "Beca de colaboración", Scholarship from the Spanish Ministry of Education for an internship in a research laboratory

2009 Scholarship from the Goethe Institut for a one-month German course in Munich, Germany

2009 Scholarship for a 1-week theoretical and practical course, CYTED (Ibero-American Program for the Development of Science and Technology) – "Inorganic Drugs and Their Mechanisms of Action", Universidad de la República, Montevideo, Uruguay

2013 Forschungskredit FK-13-090, University of Zurich

2013 Short Term Scientific Mission (STSM) funding, COST Action CM1105

Publications

J. Benítez, L. Becco, I. Correia, S. M. Leal, H. Guiset Miserachs, J. Costa Pessoa, J. Lorenzo, S. Tanco, P. Escobar, V. Moreno, D. Gambino. Vanadium polypyridyl compounds as potential antiparasitic and antitumoral agents: New achievement. *Journal of Inorganic Biochemistry*, 105, 303-312, 2011.

B. Demoro, M. Rossi, F. Caruso, D. Liebowitz, C. Olea-Azar, U. Kemmerling, J. D. Maya, H. Guiset Miserachs, V. Moreno, C. Pizzo, G. Mahler, L. Otero, D. Gambino. Potential mechanism of the anti-trypanosomal activity of organoruthenium complexes with bioactive thiosemicarbazones. *Biol Trace Elem Res.*, 153, 371-381, 2013.

H. Guiset Miserachs, M. Cipriani, J. Grau, M. Vilaseca, J. Lorenzo, A. Medeiros, M. A. Comini, D. Gambino, L. Otero, V. Moreno. Antitumor and antiparasitic activity of ruthenium compounds with polycyclic aromatic ligands. *Journal of Inorganic Biochemistry*, 150, 38-47, 2015.

H. Guiset Miserachs, D. Donghi, R. Börner, S. Johannsen, R. K. O. Sigel. Spectroscopic study on the effect of metal ions on G-quadruplex RNA folding and stability. *Journal of Biological Inorganic Chemistry*, 2016, *submitted*.

H. Guiset Miserachs, A. Marchand, D. Donghi, B. Klejevskaja, R. Börner, F. Rosu, R. Vilar, V. Gabelica, R. K. O. Sigel. Multimerization of short G4 RNA sequences as a function of metal ion concentration and flanking bases. 2016, *in preparation*.

Talks at scientific meetings

H. Guiset Miserachs, R. K. O. Sigel. "Structure and dynamics of human RNA G-quadruplex motifs", *COST CM1105 Working Group 2 Meeting*, Birmingham, UK, March 5th 2013

H. Guiset Miserachs, B. Klejevskaja, R. Vilar, R. K. O. Sigel. "Multimerization of RNA G-quadruplexes: from Zurich to London and back", *2nd International Symposium on Functional*

Metal Complexes that Bind to Biomolecules and 3rd Whole Action Meeting of the COST Action CM1105, Zurich, Switzerland, August 22nd, 2014

H. Guiset Miserachs, R. K. O. Sigel. "Effect on Metal Ions on RNA: Spectroscopic and Native Gel Studies on Human G-Quadruplexes", *7th Retreat of the CMSZH Graduate School, Zuoz*, Switzerland, January 9th, 2015

H. Guiset Miserachs, R. K. O. Sigel. "Metal ion dependency of human RNA G-quadruplexes", *Doktorandentag*, Au, Switzerland, June 4th, 2015

Selected posters at scientific meetings

H. Guiset Miserachs, L. Otero, J. Lorenzo, D. Gambino, V. Moreno, "Description of a novel ruthenium antitumor compound: [RuCl(η^6-*p*-cymene)(1,10-phenanthroline-5,6-dione][PF_6]", *11th International Symposium on Applied Bioinorganic Chemistry (ISABC11)*, Barcelona, Spain, December 2nd-5th, 2011

H. Guiset Miserachs, R. K. O. Sigel, "Role of mono- and divalent cations in the stabilization of RNA G-quadruplex structures", *11th European Bioinorganic Chemistry Conference (EUROBIC11)*, Granada, Spain, September 12th-16th, 2012

H. Guiset Miserachs, D. Donghi, R. K. O. Sigel, "Studies on the Stability and Dynamics of Human RNA G-quadruplex Structures", *4th International Meeting on G-quadruplex Nucleic Acids*, Singapore, July 1st-4th, 2013

H. Guiset Miserachs, B. Klejevskaja, R. Vilar, R. K. O. Sigel, "Study of the folding intermediates of human RNA G-quadruplexes", *1st International Symposium on Functional Metal Complexes that Bind to Biomolecules and 2nd Whole Action Meeting of the COST Action CM1105*, Barcelona, Spain, September 9th-10th, 2013.

H. Guiset Miserachs, D. Donghi, R. K. O. Sigel, "RNA G-quadruplexes: cation stabilization of NRAS and TERRA", *Meeting GDR-G-quadruplex*, Grenoble, France, March 20th-21st, 2014

H. Guiset Miserachs, B. Klejevskaja, R. Börner, D. Donghi, S. Johannsen, R. Vilar, R. K. O. Sigel, "Biologically relevant RNA G-quadruplex structure studied at the single-molecule level", *12th European Bioinorganic Chemistry Conference (EUROBIC11)*, Zurich, Switzerland, August 24th-28th, 2014

H. Guiset Miserachs, B. Klejevskaja, A. Marchand, R. Börner, D. Donghi, S. Johannsen, F. Rosu, V. Gabelica, R. Vilar, R. K. O. Sigel, "RNA G-quadruplexes stabilized by different metal ions", *Biochemistry Summer School, Villars-sur-Ollon*, Switzerland, September 1st-5th, 2014

H. Guiset Miserachs, D. Donghi, R. Börner, S. Johannsen, R. K. O. Sigel, "Unravelling biologically relevant G-quadruplexes: single-molecule studies and metal ion dependency", *5th International Meeting on G-quadruplex Nucleic Acids*, Bordeaux, France, May 26th-28th, 2015

“We know what we are, but know not what we may be.” – William Shakespeare